



SEQUENCE LISTING

<110> VERTEX PHARMACEUTICALS INC.

<120> METHODS AND COMPOSITIONS FOR RAPID DEVELOPMENT OF SCREENING ASSAYS

<130> VPI/02-143WO2

<140> PCT/US03/28078

<141> 2003-09-05

<150> 60/408,297

<151> 2002-09-05

<160> 86

<170> PatentIn version 3.2

<210> 1

<211> 795

<212> DNA

<213> Escherichia coli

<220>

<221> CDS

<222> (1)..(795)

<223> B-lactamase

<400> 1

atg	agt	cac	cca	gaa	acg	ctg	gtg	aaa	gta	aaa	gat	gct	gaa	gat	cag
Met	Ser	His	Pro	Glu	Thr	Leu	Val	Lys	Val	Lys	Asp	Ala	Glu	Asp	Gln
1									10						15

48

ttg	ggt	gca	cga	gtg	ggt	tac	atc	gaa	ctg	gat	ctc	aac	agc	ggt	aag
Leu	Gly	Ala	Arg	Val	Gly	Tyr	Ile	Glu	Leu	Asp	Leu	Asn	Ser	Gly	Lys
20								25						30	

96

atc	ctt	gag	agt	ttt	cgc	ccc	gaa	gaa	cgt	ttt	cca	atg	atg	agc	act
Ile	Leu	Glu	Ser	Phe	Arg	Pro	Glu	Glu	Arg	Phe	Pro	Met	Met	Ser	Thr
35							40				45				

144

ttt	aaa	gtt	ctg	cta	tgt	ggc	gcg	gta	tta	tcc	cgt	gtt	gac	gcc	ggg
Phe	Lys	Val	Leu	Leu	Cys	Gly	Ala	Val	Leu	Ser	Arg	Val	Asp	Ala	Gly
50							55				60				

192

caa	gag	caa	ctc	ggt	cgc	cgc	ata	cac	tat	tct	cag	aat	gac	ttg	gtt
Gln	Glu	Gln	Leu	Gly	Arg	Arg	Ile	His	Tyr	Ser	Gln	Asn	Asp	Leu	Val
65							70			75			80		

240

gag	tac	tca	cca	gtc	aca	gaa	aag	cat	ctt	acg	gat	ggc	atg	aca	gta
Glu	Tyr	Ser	Pro	Val	Thr	Glu	Lys	His	Leu	Thr	Asp	Gly	Met	Thr	Val
85								90				95			

288

aga	gaa	tta	tgc	agt	gct	gcc	ata	acc	atg	agt	gat	aac	act	gcg	gcc
Arg	Glu	Leu	Cys	Ser	Ala	Ala	Ile	Thr	Met	Ser	Asp	Asn	Thr	Ala	Ala
100								105				110			

336

aac tta ctt ctg aca acg atc gga gga ccg aag gag cta acc gct ttt Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe 115 120 125	384
ttg cac aac atg ggg gat cat gta act cgc ctt gat cgt tgg gaa ccg Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp Arg Trp Glu Pro 130 135 140	432
gag ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg cct Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro 145 150 155 160	480
gca gca atg gca aca acg ttg cgc aaa cta tta act ggc gaa cta ctt Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu 165 170 175	528
act cta gct tcc cgg caa caa tta ata gac tgg atg gag gcg gat aaa Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys 180 185 190	576
gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile 195 200 205	624
gct gat aaa tct gga gcc ggt gag cgt ggg tct cgc ggt atc att gca Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala 210 215 220	672
gca ctg ggg cca gat ggt aag ccc tcc cgt atc gta gtt atc tac acg Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr 225 230 235 240	720
acg ggg agt cag gca act atg gat gaa cga aat aga cag atc gct gag Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu 245 250 255	768
ata ggt gcc tca ctg att aag cat tgg Ile Gly Ala Ser Leu Ile Lys His Trp 260 265	795
<210> 2	
<211> 265	
<212> PRT	
<213> Escherichia coli	
<400> 2	
Met Ser His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln 1 5 10 15	
Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys 20 25 30	
Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr 35 40 45	
Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Val Asp Ala Gly 50 55 60	

Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val
 65 70 75 80
 Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val
 85 90 95
 Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala
 100 105 110
 Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe
 115 120 125
 Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp Arg Trp Glu Pro
 130 135 140
 Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro
 145 150 155 160
 Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu
 165 170 175
 Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys
 180 185 190
 Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile
 195 200 205
 Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala
 210 215 220
 Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr
 225 230 235 240
 Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu
 245 250 255
 Ile Gly Ala Ser Leu Ile Lys His Trp
 260 265

<210> 3
 <211> 858
 <212> DNA
 <213> Escherichia coli

<220>
 <221> CDS
 <222> (1)...(858)
 <223> B-lactamase

<400> 3
 atg aga att caa cat ttc cgt gtc gcc ctt att ccc ttt ttt gcg gca
 Met Arg Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
 1 5 10 15

ttt tgc ctt cct gtt ttt ggt cac cca gaa acg ctg gtg aaa gta aaa Phe Cys Leu Pro Val Phe Gly His Pro Glu Thr Leu Val Lys Val Lys 20 25 30	96
gat gct gaa gat cag ttg ggt gca cga gtg ggt tac atc gaa ctg gat Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp 35 40 45	144
ctc aac agc ggt aag atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe . 50 55 60	192
cca atg atg agc act ttt aaa gtt ctg cta tgt ggc gcg gta tta tcc Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser 65 70 75 80	240
cgt gtt gac gcc ggg caa gag caa ctc ggt cgc cgc ata cac tat tct Arg Val Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser 85 90 95	288
cag aat gac ttg gtt gag tac tca cca gtc aca gaa aag cat ctt acg Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr 100 105 110	336
gat ggc atg aca gta aga gaa tta tgc agt gct gcc ata acc atg agt Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser 115 120 125	384
gat aac act gcg gcc aac tta ctt ctg aca acg atc gga gga ccg aag Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys 130 135 140	432
gag cta acc gct ttt ttg cac aac atg ggg gat cat gta act cgc ctt Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu 145 150 155 160	480
gat cgt tgg gaa ccg gag ctg aat gaa gcc ata cca aac gac gag cgt Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg 165 170 175	528
gac acc acg atg cct gca gca atg gca aca acg ttg cgc aaa cta tta Asp Thr Thr Met Pro Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu 180 185 190	576
act ggc gaa cta ctt act cta gct tcc cg ^g caa caa tta ata gac tgg Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp 195 200 205	624
atg gag gcg gat aaa gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro 210 215 220	672
gct ggc tgg ttt att gct gat aaa tct gga gcc ggt gag cgt ggg tct Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser 225 230 235 240	720

cgc ggt atc att gca gca ctg ggg cca gat ggt aag ccc tcc cgt atc Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile 245 250 255	768
gta gtt atc tac acg acg ggg agt cag gca act atg gat gaa cga aat Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn 260 265 270	816
aga cag atc gct gag ata ggt gcc tca ctg att aag cat tgg Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp 275 280 285	858
<210> 4	
<211> 286	
<212> PRT	
<213> Escherichia coli	
<400> 4	
Met Arg Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala 1 5 10 15	
Phe Cys Leu Pro Val Phe Gly His Pro Glu Thr Leu Val Lys Val Lys 20 25 30	
Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp 35 40 45	
Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe 50 55 60	
Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser 65 70 75 80	
Arg Val Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser 85 90 95	
Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr 100 105 110	
Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser 115 120 125	
Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys 130 135 140	
Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu 145 150 155 160	
Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg 165 170 175	
Asp Thr Thr Met Pro Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu 180 185 190	
Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp 195 200 205	

Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
 210 215 220

Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
 225 230 235 240

Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
 245 250 255

Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn
 260 265 270

Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp
 275 280 285

<210> 5

<211> 795

<212> DNA

<213> Escherichia coli

<220>

<221> CDS

<222> (1)...(795)

<223> B-lactamase

<400> 5

atg ggg cac cca gaa acg ctg gtg aaa gta aaa gat gct gaa gat cag	48
Met Gly His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln	
1 5 10 15	

ttg ggt gca cga gtg ggt tac atc gaa ctg gat ctc aac agc ggt aag	96
Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys	
20 25 30	

atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc act	144
Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr	
35 40 45	

ttt aaa gtt ctg cta tgt ggc gcg gta tta tcc cgt gat gac gcc ggg	192
Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Asp Asp Ala Gly	
50 55 60	

caa gag caa ctc ggt cgc cgc ata cac tat tct cag aat gac ttg gtt	240
Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val	
65 70 75 80	

gag tac tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta	288
Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val	
85 90 95	

aga gaa tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc	336
Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala	
100 105 110	

aac tta ctt ctg aca acg atc gga gga ccg aag gag cta acc gct ttt	384
Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe	
115 120 125	

ttg cac aac atg ggg gat cat gta act cgc ctt gat cat tgg gaa ccg Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro 130 135 140	432
gag ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg cct Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro 145 150 155 160	480
gta gca atg gca aca acg ttg cgc aaa cta tta act ggc gaa cta ctt Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu 165 170 175	528
act cta gct tcc cgg caa caa tta ata gac tgg atg gag gcg gat aaa Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys 180 185 190	576
gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile 195 200 205	624
gct gat aaa tct gga gcc ggt gag cgt ggg tct cgc ggt atc att gca Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala 210 215 220	672
gca ctg ggg cca gat ggt aag ccc tcc cgt atc gta gtt atc tac acg Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr 225 230 235 240	720
acg ggg agt cag gca act atg gat gaa cga aat aga cag atc gct gag Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu 245 250 255	768
ata ggt gcc tca ctg att aag cat tgg Ile Gly Ala Ser Leu Ile Lys His Trp 260 265	795
<210> 6	
<211> 265	
<212> PRT	
<213> Escherichia coli	
<400> 6	
Met Gly His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln 1 5 10 15	
Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys 20 25 30	
Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr 35 40 45	
Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Asp Asp Ala Gly 50 55 60	
Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val 65 70 75 80	

Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val
 85 90 95

 Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala
 100 105 110

 Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe
 115 120 125

 Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro
 130 135 140

 Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro
 145 150 155 160

 Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu
 165 170 175

 Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys
 180 185 190

 Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile
 195 200 205

 Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala
 210 215 220

 Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr
 225 230 235 240

 Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu
 245 250 255

 Ile Gly Ala Ser Leu Ile Lys His Trp
 260 265

<210> 7
 <211> 792
 <212> DNA
 <213> Escherichia coli

<220>
 <221> CDS
 <222> (1)..(792)
 <223> B-lactamase

<400> 7
 atg gac cca gaa acg ctg gtg aaa gta aaa gat gct gaa gat cag ttg 48
 Met Asp Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu
 1 5 10 15

 ggt gca cga gtg ggt tac atc gaa ctg gat ctc aac agc ggt aag atc 96
 Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile
 20 25 30

ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc act ttt Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe 35 40 45	144
aaa gtt ctg cta tgt ggc gcg gta tta tcc cgt att gac gcc ggg caa Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Ile Asp Ala Gly Gln 50 55 60	192
gag caa ctc ggt cgc cgc ata cac tat tct cag aat gac ttg gtt gag Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu 65 70 75 80	240
tac tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta aga Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val Arg 85 90 95	288
gaa tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc aac Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn 100 105 110	336
tta ctt ctg aca acg atc gga gga ccg aag gag cta acc gct ttt ttg Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu 115 120 125	384
cac aac atg ggg gat cat gta act cgc ctt gat cat tgg gaa ccg gag His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro Glu 130 135 140	432
ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg cct gta Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val 145 150 155 160	480
gca atg gca aca acg ttg cgc aaa cta tta act ggc gaa cta ctt act Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr 165 170 175	528
cta gct tcc cggcaa caa tta ata gac tgg atg gag gcg gat aaa gtt Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val 180 185 190	576
gca gga cca ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att gct Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala 195 200 205	624
gat aaa tct gga gcc ggt gag cgt ggg tct cgc ggt atc att gca gca Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala 210 215 220	672
ctg ggg cca gat ggt aag ccc tcc cgt atc gta gtt atc tac acg acg Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr 225 230 235 240	720
ggg agt cag gca act atg gat gaa cga aat aga cag atc gct gag ata Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile 245 250 255	768

ggt gcc tca ctg att aag cat tgg 792
 Gly Ala Ser Leu Ile Lys His Trp
 260

 <210> 8
 <211> 264
 <212> PRT
 <213> Escherichia coli

 <400> 8
 Met Asp Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu
 1 5 10 15

 Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile
 20 25 30

 Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe
 35 40 45

 Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Ile Asp Ala Gly Gln
 50 55 60

 Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu
 65 70 75 80

 Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val Arg
 85 90 95

 Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn
 100 105 110

 Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu
 115 120 125

 His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro Glu
 130 135 140

 Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val
 145 150 155 160

 Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr
 165 170 175

 Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val
 180 185 190

 Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala
 195 200 205

 Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala
 210 215 220

 Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr
 225 230 235 240

 Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile
 245 250 255

Gly Ala Ser Leu Ile Lys His Trp
260

```
<210> 9  
<211> 786  
<212> DNA  
<213> Escherichia coli
```

<220>
<221> CDS
<222> (1)..(786)
<223> B-lactamase

gcc ttt gct ctt gaa gat aaa ctt cca agt gaa aaa cgc gag ctt tta Ala Phe Ala Leu Glu Asp Lys Leu Pro Ser Glu Lys Arg Glu Leu Leu 165 170 175	528
atc gat tgg atg aaa cga aat acc act gga gac gcc tta atc cgt gcc Ile Asp Trp Met Lys Arg Asn Thr Thr Gly Asp Ala Leu Ile Arg Ala 180 185 190	576
gga gcg gca tca tat gga acc cgg aat gac att gcc atc att tgg ccg Gly Ala Ala Ser Tyr Gly Thr Arg Asn Asp Ile Ala Ile Ile Trp Pro 195 200 205	624
cca aaa gga gat cct gtc ggt gtg ccg gac ggt tgg gaa gtg gct gat Pro Lys Gly Asp Pro Val Gly Val Pro Asp Gly Trp Glu Val Ala Asp 210 215 220	672
aaa act gtt ctt gca gta tta tcc agc agg gat aaa aag gac gcc aag Lys Thr Val Leu Ala Val Leu Ser Ser Arg Asp Lys Lys Asp Ala Lys 225 230 235 240	720
tat gat gat aaa ctt att gca gag gca aca aag gtg gta atg aaa gcc Tyr Asp Asp Lys Leu Ile Ala Glu Ala Thr Lys Val Val Met Lys Ala 245 250 255	768
tta aac atg aac ggc aaa Leu Asn Met Asn Gly Lys 260	786
<210> 10	
<211> 262	
<212> PRT	
<213> Escherichia coli)
<400> 10	
Met Lys Asp Asp Phe Ala Lys Leu Glu Glu Gln Phe Asp Ala Lys Leu 1 5 10 15	
Gly Ile Phe Ala Leu Asp Thr Gly Thr Asn Arg Thr Val Ala Tyr Arg 20 25 30	
Pro Asp Glu Arg Phe Ala Phe Ala Ser Thr Ile Lys Ala Leu Thr Val 35 40 45	
Gly Val Leu Leu Gln Gln Lys Ser Ile Glu Asp Leu Asn Gln Arg Ile 50 55 60	
Thr Tyr Thr Arg Asp Asp Leu Val Asn Tyr Asn Pro Ile Thr Glu Lys 65 70 75 80	
His Val Asp Thr Gly Met Thr Leu Lys Glu Leu Ala Asp Ala Ser Leu 85 90 95	
Arg Tyr Ser Asp Asn Ala Ala Gln Asn Leu Ile Leu Lys Gln Ile Gly 100 105 110	
Gly Pro Glu Ser Leu Lys Lys Glu Leu Arg Lys Ile Gly Asp Glu Val 115 120 125	

<210> 11
<211> 720
<212> DNA
<213> Aequorea victoria

<220>
<221> CDS
<222> (1)..(720)
<223> mutant green fluorescent protein

```

<400> 11
atg gtg agc aag ggc gag gag ctg ttc acc ggg gtg gtg ccc atc ctg      48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1           5           10          15

gtc gag ctg gac ggc gac gta aac ggc cac aag ttc agc gtg tcc ggc      96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
20          25          30

gag ggc gag ggc gat gcc acc tac ggc aag ctg acc ctg aag ttc atc      144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
35          40          45

tgc acc acc ggc aag ctg ccc gtg ccc tgg ccc acc ctc gtg acc acc      192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
50          55          60

```

ttc tcc tac ggc gtg cag tgc ttc agc cgc tac ccc gac cac atg aag Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys 65 70 75 80	240
cag cac gac ttc ttc aag tcc gcc atg ccc gaa ggc tac gtc cag gag Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu 85 90 95	288
cgc acc atc ttc ttc aag gac gac ggc aac tac aag acc cgc gcc gag Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu 100 105 110	336
gtg aag ttc gag ggc gac acc ctg gtg aac cgc atc gag ctg aag ggc Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly 115 120 125	384
atc gac ttc aag gag gac ggc aac atc ctg ggg cac aac ctg gag tac Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Asn Leu Glu Tyr 130 135 140	432
aac tac aac agc cac aac gtc tat atc atg gcc gac aag cag aag aac Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn 145 150 155 160	480
ggc atc aag gtg aac ttc aag atc cgc cac aac atc gag gac ggc agc Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser 165 170 175	528
gtg cag ctc gcc gac cac tac cag cag aac acc ccc atc ggc gac ggc Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly 180 185 190	576
ccc gtg ctg ctg ccc gac aac cac tac ctg agc acc cag tcc gcc ctg Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu 195 200 205	624
agc aaa gac ccc aac gag aag cgc gat cac atg gtc ctg ctg gag ttc Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe 210 215 220	672
gtg acc gcc gcc ggg atc act ctc ggc atg gac gag ctg tac aag taa Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys 225 230 235	720
<210> 12	
<211> 239	
<212> PRT	
<213> Aequorea victoria	
<400> 12	
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu 1 5 10 15	
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly 20 25 30	

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
 35 40 45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
 50 55 60

Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
 65 70 75 80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
 85 90 95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
 100 105 110

Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
 115 120 125

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Asn Leu Glu Tyr
 130 135 140

Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
 145 150 155 160

Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
 165 170 175

Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
 180 185 190

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
 195 200 205

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
 210 215 220

Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
 225 230 235

<210> 13

<211> 690

<212> DNA

<213> Anemonia majano

<220>

<221> CDS

<222> (1)..(690)

<223> fluorescent protein

<400> 13

atg gct ctt tca aac aag ttt atc gga gat gac atg aaa atg acc tac
 Met Ala Leu Ser Asn Lys Phe Ile Gly Asp Asp Met Lys Met Thr Tyr
 1 5 10 15

cat atg gat ggc tgt gtc aat ggg cat tac ttt acc gtc aaa ggt gaa
 His Met Asp Gly Cys Val Asn Gly His Tyr Phe Thr Val Lys Gly Glu
 20 25 30

48 96

ggc aac ggg aag cca tac gaa ggg acg cag act tcg act ttt aaa gtc Gly Asn Gly Lys Pro Tyr Glu Gly Thr Gln Thr Ser Thr Phe Lys Val 35 40 45	144
acc atg gcc aac ggt ggg ccc ctt gca ttc tcc ttt gac ata cta tct Thr Met Ala Asn Gly Gly Pro Leu Ala Phe Ser Phe Asp Ile Leu Ser 50 55 60	192
aca gtg ttc aaa tat gga aat cga tgc ttt act gcg tat cct acc agt Thr Val Phe Lys Tyr Gly Asn Arg Cys Phe Thr Ala Tyr Pro Thr Ser 65 70 75 80	240
atg ccc gac tat ttc aaa caa gca ttt cct gac gga atg tca tat gaa Met Pro Asp Tyr Phe Lys Gln Ala Phe Pro Asp Gly Met Ser Tyr Glu 85 90 95	288
agg act ttt acc tat gaa gat gga gga gtt gct aca gcc agt tgg gaa Arg Thr Phe Thr Tyr Glu Asp Gly Gly Val Ala Thr Ala Ser Trp Glu 100 105 110	336
ata agc ctt aaa ggc aac tgc ttt gag cac aaa tcc acg ttt cat gga Ile Ser Leu Lys Gly Asn Cys Phe Glu His Lys Ser Thr Phe His Gly 115 120 125	384
gtg aac ttt cct gct gat gga cct gtg atg gcg aag aag aca act ggt Val Asn Phe Pro Ala Asp Gly Pro Val Met Ala Lys Lys Thr Thr Gly 130 135 140	432
tgg gac cca tct ttt gag aaa atg act gtc tgc gat gga ata ttg aag Trp Asp Pro Ser Phe Glu Lys Met Thr Val Cys Asp Gly Ile Leu Lys 145 150 155 160	480
ggg gat gtc acc gcg ttc ctc atg ctg caa gga ggt ggc aat tac aga Gly Asp Val Thr Ala Phe Leu Met Leu Gln Gly Gly Asn Tyr Arg 165 170 175	528
tgc caa ttc cac act tct tac aag aca aaa aaa ccg gtg acg atg cca Cys Gln Phe His Thr Ser Tyr Lys Thr Lys Lys Pro Val Thr Met Pro 180 185 190	576
cca aac cat gtg gtg gaa cat cgc att gcg agg acc gac ctt gac aaa Pro Asn His Val Val Glu His Arg Ile Ala Arg Thr Asp Leu Asp Lys 195 200 205	624
ggg ggc aac agt gtt cag ctg acg gag cac gct gtt gca cat ata acc Gly Gly Asn Ser Val Gln Leu Thr Glu His Ala Val Ala His Ile Thr 210 215 220	672
tct gtt gtc cct ttc tga Ser Val Val Pro Phe 225	690

<210> 14
<211> 229
<212> PRT
<213> Anemonia majano

<400> 14
Met Ala Leu Ser Asn Lys Phe Ile Gly Asp Asp Met Lys Met Thr Tyr
1 5 10 15
His Met Asp Gly Cys Val Asn Gly His Tyr Phe Thr Val Lys Gly Glu
20 25 30
Gly Asn Gly Lys Pro Tyr Glu Gly Thr Gln Thr Ser Thr Phe Lys Val
35 40 45
Thr Met Ala Asn Gly Gly Pro Leu Ala Phe Ser Phe Asp Ile Leu Ser
50 55 60
Thr Val Phe Lys Tyr Gly Asn Arg Cys Phe Thr Ala Tyr Pro Thr Ser
65 70 75 80
Met Pro Asp Tyr Phe Lys Gln Ala Phe Pro Asp Gly Met Ser Tyr Glu
85 90 95
Arg Thr Phe Thr Tyr Glu Asp Gly Gly Val Ala Thr Ala Ser Trp Glu
100 105 110
Ile Ser Leu Lys Gly Asn Cys Phe Glu His Lys Ser Thr Phe His Gly
115 120 125
Val Asn Phe Pro Ala Asp Gly Pro Val Met Ala Lys Lys Thr Thr Gly
130 135 140
Trp Asp Pro Ser Phe Glu Lys Met Thr Val Cys Asp Gly Ile Leu Lys
145 150 155 160
Gly Asp Val Thr Ala Phe Leu Met Leu Gln Gly Gly Asn Tyr Arg
165 170 175
Cys Gln Phe His Thr Ser Tyr Lys Thr Lys Lys Pro Val Thr Met Pro
180 185 190
Pro Asn His Val Val Glu His Arg Ile Ala Arg Thr Asp Leu Asp Lys
195 200 205
Gly Gly Asn Ser Val Gln Leu Thr Glu His Ala Val Ala His Ile Thr
210 215 220
Ser Val Val Pro Phe
225

<210> 15
<211> 696
<212> DNA
<213> Zoanthus sp.

<220>
<221> CDS
<222> (1)..(696)
<223> fluorescent protein

<400> 15			
atg gct cag tca aag cac ggt cta aca aaa gaa atg aca atg aaa tac			48
Met Ala Gln Ser Lys His Gly Leu Thr Lys Glu Met Thr Met Lys Tyr			
1 5 10 15			
cgt atg gaa ggg tgc gtc gat gga cat aaa ttt gtg atc acg gga gag			96
Arg Met Glu Gly Cys Val Asp Gly His Lys Phe Val Ile Thr Gly Glu			
20 25 30			
ggc att gga tat ccg ttc aaa ggg aaa cag gct att aat ctg tgt gtg			144
Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Ala Ile Asn Leu Cys Val			
35 40 45			
gtc gaa ggt gga cca ttg cca ttt gcc gaa gac ata ttg tca gct gcc			192
Val Glu Gly Pro Leu Pro Phe Ala Glu Asp Ile Leu Ser Ala Ala			
50 55 60			
ttt aac tac gga aac agg gtt ttc act gaa tat cct caa gac ata gtt			240
Phe Asn Tyr Gly Asn Arg Val Phe Thr Glu Tyr Pro Gln Asp Ile Val			
65 70 75 80			
gac tat ttc aag aac tcg tgt cct gct gga tat aca tgg gac agg tct			288
Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Asp Arg Ser			
85 90 95			
ttt ctc ttt gag gat gga gca gtt tgc ata tgt aat gca gat ata aca			336
Phe Leu Phe Asp Gly Ala Val Cys Ile Cys Asn Ala Asp Ile Thr			
100 105 110			
gtg agt gtt gaa gaa aac tgc atg tat cat gag tcc aaa ttt tat gga			384
Val Ser Val Glu Glu Asn Cys Met Tyr His Glu Ser Lys Phe Tyr Gly			
115 120 125			
gtg aat ttt cct gct gat gga cct gtg atg aaa aag atg aca gat aac			432
Val Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Asp Asn			
130 135 140			
tgg gag cca tcc tgc gag aag atc ata cca gta cct aag cag ggg ata			480
Trp Glu Pro Ser Cys Glu Lys Ile Ile Pro Val Pro Lys Gln Gly Ile			
145 150 155 160			
ttg aaa ggg gat gtc tcc atg tac ctc ctt ctg aag gat ggt ggg cgt			528
Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg			
165 170 175			
tta cgg tgc caa ttc gac aca gtt tac aaa gca aag tct gtg cca aga			576
Leu Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Arg			
180 185 190			
aag atg ccg gac tgg cac ttc atc cag cat aag ctc acc cgt gaa gac			624
Lys Met Pro Asp Trp His Phe Ile Gln His Lys Leu Thr Arg Glu Asp			
195 200 205			
cgc agc gat gct aag aat cag aaa tgg cat ctg aca gaa cat gct att			672
Arg Ser Asp Ala Lys Asn Gln Lys Trp His Leu Thr Glu His Ala Ile			
210 215 220			

gca tcc gga tct gca ttg ccc tga
 Ala Ser Gly Ser Ala Leu Pro
 225 230

696

<210> 16
<211> 231
<212> PRT
<213> Zoanthus sp.

<400> 16
 Met Ala Gln Ser Lys His Gly Leu Thr Lys Glu Met Thr Met Lys Tyr
 1 5 10 15
 Arg Met Glu Gly Cys Val Asp Gly His Lys Phe Val Ile Thr Gly Glu
 20 25 30
 Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Ala Ile Asn Leu Cys Val
 35 40 45
 Val Glu Gly Gly Pro Leu Pro Phe Ala Glu Asp Ile Leu Ser Ala Ala
 50 55 60
 Phe Asn Tyr Gly Asn Arg Val Phe Thr Glu Tyr Pro Gln Asp Ile Val
 65 70 75 80
 Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Asp Arg Ser
 85 90 95
 Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Ala Asp Ile Thr
 100 105 110
 Val Ser Val Glu Glu Asn Cys Met Tyr His Glu Ser Lys Phe Tyr Gly
 115 120 125
 Val Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Asp Asn
 130 135 140
 Trp Glu Pro Ser Cys Glu Lys Ile Ile Pro Val Pro Lys Gln Gly Ile
 145 150 155 160
 Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Lys Asp Gly Gly Arg
 165 170 175
 Leu Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Arg
 180 185 190
 Lys Met Pro Asp Trp His Phe Ile Gln His Lys Leu Thr Arg Glu Asp
 195 200 205
 Arg Ser Asp Ala Lys Asn Gln Lys Trp His Leu Thr Glu His Ala Ile
 210 215 220
 Ala Ser Gly Ser Ala Leu Pro
 225 230

<210> 17
 <211> 696
 <212> DNA
 <213> Zoanthus sp.

 <220>
 <221> CDS
 <222> (1)..(696)
 <223> fluorescent protein

 <400> 17

atg gct cat tca aag cac ggt cta aaa gaa gaa atg aca atg aaa tac	48
Met Ala His Ser Lys His Gly Leu Lys Glu Glu Met Thr Met Lys Tyr	
1 5 10 15	
 cac atg gaa ggg tgc gtc aac gga cat aaa ttt gtg atc acg ggc gaa	96
His Met Glu Gly Cys Val Asn Gly His Lys Phe Val Ile Thr Gly Glu	
20 25 30	
 ggc att gga tat ccg ttc aaa ggg aaa cag act att aat ctg tgt gtg	144
Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Thr Ile Asn Leu Cys Val	
35 40 45	
 atc gaa ggg gga cca ttg cca ttt tcc gaa gac ata ttg tca gct ggc	192
Ile Glu Gly Pro Leu Pro Phe Ser Glu Asp Ile Leu Ser Ala Gly	
50 55 60	
 ttt aag tac gga gac agg att ttc act gaa tat cct caa gac ata gta	240
Phe Lys Tyr Gly Asp Arg Ile Phe Thr Glu Tyr Pro Gln Asp Ile Val	
65 70 75 80	
 gac tat ttc aag aac tcg tgt cct gct gga tat aca tgg ggc agg tct	288
Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Gly Arg Ser	
85 90 95	
 ttt ctc ttt gag gat gga gca gtc tgc ata tgc aat gta gat ata aca	336
Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Val Asp Ile Thr	
100 105 110	
 gtg agt gtc aaa gaa aac tgc att tat cat aag agc ata ttt aat gga	384
Val Ser Val Lys Glu Asn Cys Ile Tyr His Lys Ser Ile Phe Asn Gly	
115 120 125	
 atg aat ttt cct gct gat gga cct gtg atg aaa aag atg aca act aac	432
Met Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Thr Asn	
130 135 140	
 tgg gaa gca tcc tgc gag aag atc atg cca gta cct aag cag ggg ata	480
Trp Glu Ala Ser Cys Glu Lys Ile Met Pro Val Pro Lys Gln Gly Ile	
145 150 155 160	
 ctg aaa ggg gat gtc tcc atg tac ctc ctt ctg aag gat ggt ggg cgt	528
Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Lys Asp Gly Gly Arg	
165 170 175	
 tac cgg tgc cag ttc gac aca gtt tac aaa gca aag tct gtg cca agt	576
Tyr Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Ser	
180 185 190	

aag atg ccg gag tgg cac ttc atc cag cat aag ctc ctc cgt gaa gac Lys Met Pro Glu Trp His Phe Ile Gln His Lys Leu Leu Arg Glu Asp 195 200 205	624
cgc agc gat gct aag aat cag aag tgg cag ctg aca gag cat gct att Arg Ser Asp Ala Lys Asn Gln Lys Trp Gln Leu Thr Glu His Ala Ile 210 215 220	672
gca ttc cct tct gcc ttg gcc tga Ala Phe Pro Ser Ala Leu Ala 225 230	696
<210> 18	
<211> 231	
<212> PRT	
<213> Zoanthus sp.	
<400> 18	
Met Ala His Ser Lys His Gly Leu Lys Glu Glu Met Thr Met Lys Tyr 1 5 10 15	
His Met Glu Gly Cys Val Asn Gly His Lys Phe Val Ile Thr Gly Glu 20 25 30	
Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Thr Ile Asn Leu Cys Val 35 40 45	
Ile Glu Gly Gly Pro Leu Pro Phe Ser Glu Asp Ile Leu Ser Ala Gly 50 55 60	
Phe Lys Tyr Gly Asp Arg Ile Phe Thr Glu Tyr Pro Gln Asp Ile Val 65 70 75 80	
Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Gly Arg Ser 85 90 95	
Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Val Asp Ile Thr 100 105 110	
Val Ser Val Lys Glu Asn Cys Ile Tyr His Lys Ser Ile Phe Asn Gly 115 120 125	
Met Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Thr Asn 130 135 140	
Trp Glu Ala Ser Cys Glu Lys Ile Met Pro Val Pro Lys Gln Gly Ile 145 150 155 160	
Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg 165 170 175	
Tyr Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Ser 180 185 190	
Lys Met Pro Glu Trp His Phe Ile Gln His Lys Leu Leu Arg Glu Asp 195 200 205	

Arg Ser Asp Ala Lys Asn Gln Lys Trp Gln Leu Thr Glu His Ala Ile
 210 215 220

Ala Phe Pro Ser Ala Leu Ala
225 230

<210> 19
<211> 699
<212> DNA
<213> *Discosoma striata*

<220>
<221> CDS
<222> (1)..(699)
<223> fluorescent protein

```

<400> 19
atg agt tgt tcc aag agt gtg atc aag gaa gaa atg ttg atc gat ctt      48
Met Ser Cys Ser Lys Ser Val Ile Lys Glu Glu Met Leu Ile Asp Leu
1           5           10          15


```

cat ctg gaa gga acg ttc aat ggg cac tac ttt gaa ata aaa ggc aaa
 His Leu Glu Gly Thr Phe Asn Gly His Tyr Phe Glu Ile Lys Gly Lys
 20 25 30

gga aaa gga cag cct aat gaa ggc acc aat acc gtc acg ctc gag gtt 144
Gly Lys Gly Gln Pro Asn Glu Gly Thr Asn Thr Val Thr Leu Glu Val .
35 40 45

acc aag ggt gga cct ctg cca ttt ggt tgg cat att ttg tgc cca caa
 Thr Lys Gly Gly Pro Leu Pro Phe Gly Trp His Ile Leu Cys Pro Gln
 50 55 60

ttt cag tat gga aac aag gca ttt gtc cac cac cct gac aac ata cat
 Phe Gln Tyr Gly Asn Lys Ala Phe Val His His Pro Asp Asn Ile His
 65 70 75 80

gat tat cta aag ctg tca ttt ccg gag gga tat aca tgg gaa cg^g tcc 288
 Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ser
 85 90 95

atg cac ttt gaa gac ggt ggc ttg tgt tgt atc acc aat gat atc agt 336
 Met His Phe Glu Asp Gly Gly Leu Cys Cys Ile Thr Asn Asp Ile Ser
 100 105 110

ttg aca ggc aac tgt ttc tac tac gac atc aag ttc act ggc ttg aac 384
 Leu Thr Gly Asn Cys Phe Tyr Tyr Asp Ile Lys Phe Thr Gly Leu Asn
 115 120 125

ttt cct cca aat gga ccc gtt gtg cag aag aag aca act ggc tgg gaa 432
 Phe Pro Pro Asn Gly Pro Val Val Gln Lys Lys Thr Thr Gly Trp Glu
 130 135 140

ccg agc act gag cgt ttg tat cct cgt gat ggt gtg ttg ata gga gac
 Pro Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Ile Gly Asp
 145 150 155 160

atc cat cat gct ctg aca gtt gaa gga ggt ggt cat tac gca tgt gac Ile His His Ala Leu Thr Val Glu Gly Gly His Tyr Ala Cys Asp 165 170 175	528
att aaa act gtt tac agg gcc aag aag gcc gcc ttg aag atg cca ggg Ile Lys Thr Val Tyr Arg Ala Lys Lys Ala Ala Leu Lys Met Pro Gly 180 185 190	576
tat cac tat gtt gac acc aaa ctg gtt ata tgg aac aac gac aaa gaa Tyr His Tyr Val Asp Thr Lys Leu Val Ile Trp Asn Asn Asp Lys Glu 195 200 205	624
ttc atg aaa gtt gag gag cat gaa atc gcc gtt gca cgc cac cat ccg Phe Met Lys Val Glu Glu His Glu Ile Ala Val Ala Arg His His Pro 210 215 220	672
ttc tat gag cca aag aag gat aag taa Phe Tyr Glu Pro Lys Lys Asp Lys 225 230	699
<210> 20	
<211> 232	
<212> PRT	
<213> Discosoma striata	
<400> 20	
Met Ser Cys Ser Lys Ser Val Ile Lys Glu Glu Met Leu Ile Asp Leu 1 5 10 15	
His Leu Glu Gly Thr Phe Asn Gly His Tyr Phe Glu Ile Lys Gly Lys 20 25 30	
Gly Lys Gly Gln Pro Asn Glu Gly Thr Asn Thr Val Thr Leu Glu Val 35 40 45	
Thr Lys Gly Gly Pro Leu Pro Phe Gly Trp His Ile Leu Cys Pro Gln 50 55 60	
Phe Gln Tyr Gly Asn Lys Ala Phe Val His His Pro Asp Asn Ile His 65 70 75 80	
Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ser 85 90 95	
Met His Phe Glu Asp Gly Gly Leu Cys Cys Ile Thr Asn Asp Ile Ser 100 105 110	
Leu Thr Gly Asn Cys Phe Tyr Tyr Asp Ile Lys Phe Thr Gly Leu Asn 115 120 125	
Phe Pro Pro Asn Gly Pro Val Val Gln Lys Lys Thr Thr Gly Trp Glu 130 135 140	
Pro Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Ile Gly Asp 145 150 155 160	

Ile His His Ala Leu Thr Val Glu Gly Gly Gly His Tyr Ala Cys Asp
165 170 175

Ile Lys Thr Val Tyr Arg Ala Lys Lys Ala Ala Leu Lys Met Pro Gly
180 185 190

Tyr His Tyr Val Asp Thr Lys Leu Val Ile Trp Asn Asn Asp Lys Glu
 195 200 205

Phe Met Lys Val Glu Glu His Glu Ile Ala Val Ala Arg His His Pro
210 215 220

Phe Tyr Glu Pro Lys Lys Asp Lys
225 230

<210> 21
<211> 678
<212> DNA
<213> *Discosoma* sp.

```
<220>
<221> CDS
<222> (1)..(678)
<223> fluorescent protein
```

<400> 21
atg agg tct tcc aag aat gtt atc aag gag ttc atg agg ttt aag gtt 48
Met Arg Ser Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val
1 5 10 15

cgc atg gaa gga acg gtc aat ggg cac gag ttt gaa ata gaa ggc gaa 96
 Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu
 20 25 30

gga gag ggg agg cca tac gaa ggc cac aat acc gta aag ctt aag gta 144
 Gly Glu Gly Arg Pro Tyr Glu Gly His Asn Thr Val Lys Leu Lys Val
 35 40 45

acc aag ggg gga cct ttg cca ttt gct tgg gat att ttg tca cca caa 192
 Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln
 50 55 60

ttt cag tat gga agc aag gta tat gtc aag cac cct gcc gac ata cca
 Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro
 65 70 75 80

gac tat aaa aag ctg tca ttt cct gaa gga ttt aaa tgg gaa agg gtc 288
 Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val
 85 90 95

atg aac ttt gaa gac ggt ggc gtc gtt act gta acc cag gat tcc agt 336
 Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser
 100 105 110

ttg cag gat ggc tgt ttc atc tac aag gtc aag ttc att ggc gtg aac 384
 Leu Gln Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val Asn
 115 120 125

ttt cct tcc gat gga cct gtt atg caa aag aac atg ggc tgg gaa Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu 130 135 140	432
gcc agc act gag cgt ttg tat cct cgt gat ggc gtg ttg aaa gga gag Ala Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu 145 150 155 160	480
att cat aag gct ctg aag ctg aaa gac ggt ggt cat tac cta gtt gaa Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val Glu 165 170 175	528
ttc aaa agt att tac atg gca aag aag cct gtg cag cta cca ggg tac Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Tyr 180 185 190	576
tac tat gtt gac tcc aaa ctg gat ata aca agc cac aac gaa gac tat Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr 195 200 205	624
aca atc gtt gag cag tat gaa aga acc gag gga cgc cac cat ctg ttc Thr Ile Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu Phe 210 215 220	672
ctt taa Leu 225	678
<210> 22	
<211> 225	
<212> PRT	
<213> Discosoma sp.	
<400> 22	
Met Arg Ser Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val 1 5 10 15	
Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu 20 25 30	
Gly Glu Gly Arg Pro Tyr Glu Gly His Asn Thr Val Lys Leu Lys Val 35 40 45	
Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln 50 55 60	
Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro 65 70 75 80	
Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val 85 90 95	
Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser 100 105 110	

26/102

Leu Gln Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val Asn
115 120 125

Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu
130 135 140

Ala Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu
145 150 155 160

Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val Glu
165 170 175

Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Tyr
 180 185 190

Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr
195 200 205

Thr Ile Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu Phe
210 215 220

Leu
225

<210> 23
<211> 801
<212> DNA
<213> Clavularia sp.

```
<220>
<221> CDS
<222> (1)..(801)
<223> fluorescent protein
```

```

<400> 23
atg aag tgt aaa ttt gtg ttc tgc ctg tcc ttc ttg gtc ctc gcc atc      48
Met Lys Cys Lys Phe Val Phe Cys Leu Ser Phe Leu Val Leu Ala Ile
1          5           10          15

```

aca aac gcg aac att ttt ttg aga aac gag gct gac tta gaa gag aag 96
 Thr Asn Ala Asn Ile Phe Leu Arg Asn Glu Ala Asp Leu Glu Glu Lys
 20 25 30

aca ttg aga ata cca aaa gct cta acc acc atg ggt gtg att aaa cca 144
 Thr Leu Arg Ile Pro Lys Ala Leu Thr Thr Met Gly Val Ile Lys Pro
 35 40 45

gac atg aag att aag ctg aag atg gaa gga aat gta aac ggg cat gct 192
 Asp Met Lys Ile Lys Leu Lys Met Glu Gly Asn Val Asn Gly His Ala
 50 55 60

ttt gtg atc gaa gga gaa gga gaa gga aag cct tac gat ggg aca cac
 Phe Val Ile Glu Gly Glu Gly Glu Gly Lys Pro Tyr Asp Gly Thr His
 65 70 75 80

act tta aac ctg gaa gtg aag gaa ggt gcg cct ctg cct ttt tct tac Thr Leu Asn Leu Glu Val Lys Glu Gly Ala Pro Leu Pro Phe Ser Tyr 85 90 95	288
gat atc ttg tca aac gcg ttc cag tac gga aac aga gca ttg aca aaa Asp Ile Leu Ser Asn Ala Phe Gln Tyr Gly Asn Arg Ala Leu Thr Lys 100 105 110	336
tac cca gac gat ata gca gac tat ttc aag cag tcg ttt ccc gag gga Tyr Pro Asp Asp Ile Ala Asp Tyr Phe Lys Gln Ser Phe Pro Glu Gly 115 120 125	384
tat tcc tgg gaa aga acc atg act ttt gaa gac aaa ggc att gtc aaa Tyr Ser Trp Glu Arg Thr Met Thr Phe Glu Asp Lys Gly Ile Val Lys 130 135 140	432
gtg aaa agt gac ata agc atg gag gaa gac tcc ttt atc tat gaa att Val Lys Ser Asp Ile Ser Met Glu Glu Asp Ser Phe Ile Tyr Glu Ile 145 150 155 160	480
cgt ttt gat ggg atg aac ttt cct ccc aat ggt ccg gtt atg cag aaa Arg Phe Asp Gly Met Asn Phe Pro Pro Asn Gly Pro Val Met Gln Lys 165 170 175	528
aaa act ttg aag tgg gaa cca tcc act gag att atg tac gtg cgt gat Lys Thr Leu Lys Trp Glu Pro Ser Thr Glu Ile Met Tyr Val Arg Asp 180 185 190	576
gga gtg ctg gtc gga gat att agc cat tct ctg ttg ctg gag gga ggt Gly Val Leu Val Gly Asp Ile Ser His Ser Leu Leu Glu Gly Gly 195 200 205	624
ggc cat tac cga tgt gac ttc aaa agt att tac aaa gca aaa aaa gtt Gly His Tyr Arg Cys Asp Phe Lys Ser Ile Tyr Lys Ala Lys Lys Val 210 215 220	672
gtc aaa ttg cca gac tat cac ttt gtg gac cat cgc att gag atc ttg Val Lys Leu Pro Asp Tyr His Phe Val Asp His Arg Ile Glu Ile Leu 225 230 235 240	720
aac cat gac aag gat tac aac aaa gta acg ctg tat gag aat gca gtt Asn His Asp Lys Asp Tyr Asn Lys Val Thr Leu Tyr Glu Asn Ala Val 245 250 255	768
gct cgc tat tct ttg ctg cca agt cag gcc tag Ala Arg Tyr Ser Leu Leu Pro Ser Gln Ala 260 265	801
<210> 24	
<211> 266	
<212> PRT	
<213> Clavularia sp.	
<400> 24	
Met Lys Cys Lys Phe Val Phe Cys Leu Ser Phe Leu Val Leu Ala Ile 1 5 10 15	

Thr Asn Ala Asn Ile Phe Leu Arg Asn Glu Ala Asp Leu Glu Glu Lys
 20 25 30

Thr Leu Arg Ile Pro Lys Ala Leu Thr Thr Met Gly Val Ile Lys Pro
 35 40 45

Asp Met Lys Ile Lys Leu Lys Met Glu Gly Asn Val Asn Gly His Ala
 50 55 60

Phe Val Ile Glu Gly Glu Gly Lys Pro Tyr Asp Gly Thr His
 65 70 75 80

Thr Leu Asn Leu Glu Val Lys Glu Gly Ala Pro Leu Pro Phe Ser Tyr
 85 90 95

Asp Ile Leu Ser Asn Ala Phe Gln Tyr Gly Asn Arg Ala Leu Thr Lys
 100 105 110

Tyr Pro Asp Asp Ile Ala Asp Tyr Phe Lys Gln Ser Phe Pro Glu Gly
 115 120 125

Tyr Ser Trp Glu Arg Thr Met Thr Phe Glu Asp Lys Gly Ile Val Lys
 130 135 140

Val Lys Ser Asp Ile Ser Met Glu Glu Asp Ser Phe Ile Tyr Glu Ile
 145 150 155 160

Arg Phe Asp Gly Met Asn Phe Pro Pro Asn Gly Pro Val Met Gln Lys
 165 170 175

Lys Thr Leu Lys Trp Glu Pro Ser Thr Glu Ile Met Tyr Val Arg Asp
 180 185 190

Gly Val Leu Val Gly Asp Ile Ser His Ser Leu Leu Leu Glu Gly Gly
 195 200 205

Gly His Tyr Arg Cys Asp Phe Lys Ser Ile Tyr Lys Ala Lys Lys Val
 210 215 220

Val Lys Leu Pro Asp Tyr His Phe Val Asp His Arg Ile Glu Ile Leu
 225 230 235 240

Asn His Asp Lys Asp Tyr Asn Lys Val Thr Leu Tyr Glu Asn Ala Val
 245 250 255

Ala Arg Tyr Ser Leu Leu Pro Ser Gln Ala
 260 265

<210> 25
 <211> 881
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 25
 Met Lys Leu Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu
 1 5 10 15

Lys Lys Leu Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu
 20 25 30

Lys Asn Asn Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro
 35 40 45

Leu Thr Arg Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu
 50 55 60

Glu Gln Leu Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile
 65 70 75 80

Leu Lys Met Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu
 85 90 95

Phe Val Gln Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala
 100 105 110

Ser Val Glu Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser
 115 120 125

Ala Thr Ser Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu
 130 135 140

Thr Val Ser Ile Asp Ser Ala Ala His His Asp Asn Ser Thr Ile Pro
 145 150 155 160

Leu Asp Phe Met Pro Arg Asp Ala Leu His Gly Phe Asp Trp Ser Glu
 165 170 175

Glu Asp Asp Met Ser Asp Gly Leu Pro Phe Leu Lys Thr Asp Pro Asn
 180 185 190

Asn Asn Gly Phe Phe Gly Asp Gly Ser Leu Leu Cys Ile Leu Arg Ser
 195 200 205

Ile Gly Phe Lys Pro Glu Asn Tyr Thr Asn Ser Asn Val Asn Arg Leu
 210 215 220

Pro Thr Met Ile Thr Asp Arg Tyr Thr Leu Ala Ser Arg Ser Thr Thr
 225 230 235 240

Ser Arg Leu Leu Gln Ser Tyr Leu Asn Asn Phe His Pro Tyr Cys Pro
 245 250 255

Ile Val His Ser Pro Thr Leu Met Met Leu Tyr Asn Asn Gln Ile Glu
 260 265 270

Ile Ala Ser Lys Asp Gln Trp Gln Ile Leu Phe Asn Cys Ile Leu Ala
 275 280 285

Ile Gly Ala Trp Cys Ile Glu Gly Glu Ser Thr Asp Ile Asp Val Phe
 290 295 300

Tyr Tyr Gln Asn Ala Lys Ser His Leu Thr Ser Lys Val Phe Glu Ser
 305 310 315 320

Gly Ser Ile Ile Leu Val Thr Ala Leu His Leu Leu Ser Arg Tyr Thr
 325 330 335
 Gln Trp Arg Gln Lys Thr Asn Thr Ser Tyr Asn Phe His Ser Phe Ser
 340 345 350
 Ile Arg Met Ala Ile Ser Leu Gly Leu Asn Arg Asp Leu Pro Ser Ser
 355 360 365
 Phe Ser Asp Ser Ser Ile Leu Glu Gln Arg Arg Arg Ile Trp Trp Ser
 370 375 380
 Val Tyr Ser Trp Glu Ile Gln Leu Ser Leu Leu Tyr Gly Arg Ser Ile
 385 390 395 400
 Gln Leu Ser Gln Asn Thr Ile Ser Phe Pro Ser Ser Val Asp Asp Val
 405 410 415
 Gln Arg Thr Thr Gly Pro Thr Ile Tyr His Gly Ile Ile Glu Thr
 420 425 430
 Ala Arg Leu Leu Gln Val Phe Thr Lys Ile Tyr Glu Leu Asp Lys Thr
 435 440 445
 Val Thr Ala Glu Lys Ser Pro Ile Cys Ala Lys Lys Cys Leu Met Ile
 450 455 460
 Cys Asn Glu Ile Glu Glu Val Ser Arg Gln Ala Pro Lys Phe Leu Gln
 465 470 475 480
 Met Asp Ile Ser Thr Thr Ala Leu Thr Asn Leu Leu Lys Glu His Pro
 485 490 495
 Trp Leu Ser Phe Thr Arg Phe Glu Leu Lys Trp Lys Gln Leu Ser Leu
 500 505 510
 Ile Ile Tyr Val Leu Arg Asp Phe Phe Thr Asn Phe Thr Gln Lys Lys
 515 520 525
 Ser Gln Leu Glu Gln Asp Gln Asn Asp His Gln Ser Tyr Glu Val Lys
 530 535 540
 Arg Cys Ser Ile Met Leu Ser Asp Ala Ala Gln Arg Thr Val Met Ser
 545 550 555 560
 Val Ser Ser Tyr Met Asp Asn His Asn Val Thr Pro Tyr Phe Ala Trp
 565 570 575
 Asn Cys Ser Tyr Tyr Leu Phe Asn Ala Val Leu Val Pro Ile Lys Thr
 580 585 590
 Leu Leu Ser Asn Ser Lys Ser Asn Ala Glu Asn Asn Glu Thr Ala Gln
 595 600 605
 Leu Leu Gln Gln Ile Asn Thr Val Leu Met Leu Leu Lys Lys Leu Ala
 610 615 620

Thr Phe Lys Ile Gln Thr Cys Glu Lys Tyr Ile Gln Val Leu Glu Glu
 625 630 635 640

Val Cys Ala Pro Phe Leu Leu Ser Gln Cys Ala Ile Pro Leu Pro His
 645 650 655

Ile Ser Tyr Asn Asn Ser Asn Gly Ser Ala Ile Lys Asn Ile Val Gly
 660 665 670

Ser Ala Thr Ile Ala Gln Tyr Pro Thr Leu Pro Glu Glu Asn Val Asn
 675 680 685

Asn Ile Ser Val Lys Tyr Val Ser Pro Gly Ser Val Gly Pro Ser Pro
 690 695 700

Val Pro Leu Lys Ser Gly Ala Ser Phe Ser Asp Leu Val Lys Leu Leu
 705 710 715 720

Ser Asn Arg Pro Pro Ser Arg Asn Ser Pro Val Thr Ile Pro Arg Ser
 725 730 735

Thr Pro Ser His Arg Ser Val Thr Pro Phe Leu Gly Gln Gln Gln
 740 745 750

Leu Gln Ser Leu Val Pro Leu Thr Pro Ser Ala Leu Phe Gly Gly Ala
 755 760 765

Asn Phe Asn Gln Ser Gly Asn Ile Ala Asp Ser Ser Leu Ser Phe Thr
 770 775 780

Phe Thr Asn Ser Ser Asn Gly Pro Asn Leu Ile Thr Thr Gln Thr Asn
 785 790 795 800

Ser Gln Ala Leu Ser Gln Pro Ile Ala Ser Ser Asn Val His Asp Asn
 805 810 815

Phe Met Asn Asn Glu Ile Thr Ala Ser Lys Ile Asp Asp Gly Asn Asn
 820 825 830

Ser Lys Pro Leu Ser Pro Gly Trp Thr Asp Gln Thr Ala Tyr Asn Ala
 835 840 845

Phe Gly Ile Thr Thr Gly Met Phe Asn Thr Thr Thr Met Asp Asp Val
 850 855 860

Tyr Asn Tyr Leu Phe Asp Asp Glu Asp Thr Pro Pro Asn Pro Lys Lys
 865 870 875 880

Glu

<210> 26
 <211> 202
 <212> PRT
 <213> *Saccharomyces cerevisiae*

32/102

<400> 26
Met Lys Ala Leu Thr Ala Arg Gln Gln Glu Val Phe Asp Leu Ile Arg
1 5 10 15
Asp His Ile Ser Gln Thr Gly Met Pro Pro Thr Arg Ala Glu Ile Ala
20 25 30
Gln Arg Leu Gly Phe Arg Ser Pro Asn Ala Ala Glu Glu His Leu Lys
35 40 45
Ala Leu Ala Arg Lys Gly Val Ile Glu Ile Val Ser Gly Ala Ser Arg
50 55 60
Gly Ile Arg Leu Leu Gln Glu Glu Glu Gly Leu Pro Leu Val Gly
65 70 75 80
Arg Val Ala Ala Gly Glu Pro Leu Leu Ala Gln Gln His Ile Glu Gly
85 90 95
His Tyr Gln Val Asp Pro Ser Leu Phe Lys Pro Asn Ala Asp Phe Leu
100 105
Leu Arg Val Ser Gly Met Ser Met Lys Asp Ile Gly Ile Met Asp Gly
115 120 125
Asp Leu Leu Ala Val His Lys Thr Gln Asp Val Arg Asn Gly Gln Val
130 135 140
Val Val Ala Arg Ile Asp Asp Glu Val Thr Val Lys Arg Leu Lys Lys
145 150 155 160
Gln Gly Asn Lys Val Glu Leu Leu Pro Glu Asn Ser Glu Phe Lys Pro
165 170 175
Ile Val Val Asp Leu Arg Gln Gln Ser Phe Thr Ile Glu Gly Leu Ala
180 185 190
Val Gly Val Ile Arg Asn Gly Asp Trp Leu
195 200

<210> 27

<211> 17

<212> DNA

<213> Unknown Organism

<220>

<223> Description of Unknown Organism: Nucleotide sequence that
is bound by GAL4

<220>

<221> modified_base

<222> (4)..(4)

<223> any purine

<220>
 <221> modified_base
 <222> (5)..(5)
 <223> any nucleotide

<220>
 <221> modified_base
 <222> (6)..(6)
 <223> any nucleotide

<220>
 <221> modified_base
 <222> (7)..(7)
 <223> any purine

<220>
 <221> modified_base
 <222> (9)..(9)
 <223> any pyrimidine

<220>
 <221> modified_base
 <222> (10)..(10)
 <223> any nucleotide

<220>
 <221> modified_base
 <222> (11)..(11)
 <223> any pyrimidine

<220>
 <221> modified_base
 <222> (12)..(12)
 <223> any nucleotide

<220>
 <221> modified_base
 <222> (14)..(14)
 <223> any nucleotide

<400> 27
 cggnnnnncnn nncnccg

17

<210> 28
 <211> 14
 <212> DNA
 <213> Unknown Organism

<220>
 <223> Description of Unknown Organism: Nucleotide sequence
 that binds LexA

<220>
 <221> modified_base
 <222> (6)..(9)
 <223> any nucleotide

<400> 28	cgaacnnnng ttcg	14
<210> 29		
<211> 25		
<212> DNA		
<213> Artificial Sequence		
<220>		
<223> Description of Artificial Sequence: Synthetic primer		
<400> 29		
ctagacgttg acattgatta ttgac		25
<210> 30		
<211> 25		
<212> DNA		
<213> Artificial Sequence		
<220>		
<223> Description of Artificial Sequence: Synthetic primer		
<400> 30		
tctagagcca gtaaggcgtg ggttc		25
<210> 31		
<211> 28		
<212> DNA		
<213> Artificial Sequence		
<220>		
<223> Description of Artificial Sequence: Synthetic primer		
<400> 31		
gcggccgccc attgcattgg gattggtc		28
<210> 32		
<211> 26		
<212> DNA		
<213> Artificial Sequence		
<220>		
<223> Description of Artificial Sequence: Synthetic primer		
<400> 32		
gcggccgccc tccaaagtctt tatctg		26
<210> 33		
<211> 28		
<212> DNA		
<213> Artificial Sequence		

```

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 33
ggtaccggct cgttagagaaa tatgaacc                                28

<210> 34
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 34
ctcgagagag actgaatttc cctttt                                26

<210> 35
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 35
gcggccgcga aggaagatag ccaaggctta                                30

<210> 36
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 36
gcggccgcctt aaagatgccaa cactgctggt ct                                32

<210> 37
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 37
gtcgacacct gtcggctact gggata                                26

<210> 38
<211> 26
<212> DNA

```

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic primer

<400> 38

gatatcggtgg gactctgttt ccagca

26

<210> 39

<211> 6795

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic recombinant
DNA vector sequence

<220>

<221> modified_base

<222> (892)..(892)

<223> unknown nucleotide

<400> 39

aagtccggca	cctcgac	gcggatttcg	gttccaacaa	tgtcctgacg	gacaatggcc	60
gcataaacaggc	ggtcattgac	tggagcgagg	cgatgttcgg	ggattcccaa	tacgagggtcg	120
ccaacatctt	cttctggagg	ccgtgggtgg	cttgtatgg	gcagcagacg	cgctacttcg	180
agcggaggca	tccggagctt	gcaggatcgc	cgcggctccg	ggcgtatatg	ctccgcattg	240
gtcttgacca	acttatacg	agcttgggtt	acggcaattt	cgatgtatgc	gcttgggcgc	300
agggtcgatg	cgacgcatac	gtccgatccg	gagccgggac	tgtcgggctgt	acacaaatcg	360
cccgagaag	cgcggccgtc	tggaccgatg	gctgtgtaga	agtactcgcc	gatagtggaa	420
accgacgccc	cagcactcgt	ccgaggggca	agaatgcct	gagaaaggaa	gtgagctgta	480
aaggctgagc	tctctctctg	acgtatgtag	cctctggta	gcttcgtcac	tcactgttct	540
tgactcagca	tggcaatctg	atgaaatccc	agctgttaagt	ctgcataaat	tgatgatcta	600
ttaaaacaata	aagatgtcca	ctaaaatgg	agttttttac	tgtcataactt	tgttaagaag	660
ggtgagaaca	gagtagctac	atttgaatg	gaaggattgg	agctacgggg	gtgggggtgg	720
gggtgggatt	agataaatgc	ctgctctta	ctgaaggctc	tttactattg	ctttatgata	780
atgtttcata	gttggatatac	ataattttaa	caagcaaaac	caaattaagg	gccagctcat	840
tcctccactc	atgatctata	gatctataga	tctctcggt	gatcattgtt	tntctgatcc	900
actggaaagct	tatcgatacc	gtcgacctcg	agggggggcc	cggtacccag	cttttgttcc	960
cttttagtgag	ggttaattgc	gchgcttggcg	taatcatgg	catagctgtt	tcctgtgtga	1020
aattgttatac	cgctcacaat	tccacacaac	atacgagccg	gaagcataaa	gtgtaaagcc	1080
tggggtgccct	aattagtgag	ctaactcaca	ttaatttgctgt	tgcgctact	gccccgtttc	1140
cagtcgggaa	acctgtcgtg	ccagctgcat	taatgaatcg	gccaacgcgc	ggggagaggc	1200
ggtttgcgt	ttggggcgctc	ttccgccttc	tgcgtcaactg	actcgctgctg	ctcggtcggt	1260
cggctgcggc	gagcgggtato	agctcactca	aaggcggtaa	tacggttatac	cacagaatca	1320
ggggataaacg	caggaaagaa	catgtgagca	aaaggccagc	aaaaggccag	gaaccgtaaa	1380
aaggccgcgt	tgtggcggtt	tttccatagg	ctccgcccc	ctgacgagca	tcacaaaaat	1440
cgacgctaa	gtcagaggtg	gcaaaaaaccg	acaggactat	aaagatacc	ggcgtttccc	1500
cctggaaagct	ccctcggtcg	cttcctgtt	ccgaccctgc	cgcttaccgg	ataacctgtcc	1560
gcctttctcc	cttcgggaag	cgtggcgctt	tctcatagct	cacgctgttag	gtatctcagt	1620
tcgggtgttag	tcgttcgtc	caagctgggc	tgtgtgcacg	aaccccccgt	tcagcccgac	1680
cgctgcgcct	tatccggtaa	ctatcgctt	gagtccaaacc	cggtaaagaca	cgacttatcg	1740
ccactggcag	cagccactgg	taacagatt	agcagagcga	ggtatgttag	cggtgttaca	1800
gagttcttga	agtgggtggcc	taactacggc	tacactagaa	ggacagtatt	tgttatctgc	1860
gctctgtga	agccagttac	cttcggaaaa	agagttggta	gctcttgatc	cgccaaacaa	1920
accaccgctg	gtacgggtgg	ttttttgtt	tgcaaggcgc	agattacgcg	cagaaaaaaaa	1980
ggatctcaag	aagatccctt	gatctttct	acggggctcg	acgctcagt	gaacgaaaac	2040

tcacgttaag ggatttttgt catgagatta tcaaaaagga tcttcaccta gatccttttta	2100
aattaaaaat gaagttttaa atcaactaa agtatatatg agtaaacttg gtctgacagt	2160
taccaatgct taatcagtga ggcacctata tcagcgatct gtctatttcg ttcatccata	2220
gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc atctggcccc	2280
agtgtctcaa tgataccgcg agacccacgc tcaccggctc cagatttatac agcaataaaac	2340
cagccagccg gaaggggccga gcgcagaagt ggtcctgc当地 ctttatccgc ctccatccag	2400
tctattaaatt gttgccggga agctagaga agtagttcgc cagttaatag ttgcgc当地	2460
gttggccca ttgctacagg catgtgggtc tcacgtcgt cggttggat ggettcatcc	2520
agctccggg cccaacgatc aaggcgagg acatgatccc ccatgttgc当地 caaaaaaaaagcg	2580
gttagctct tcggccctcc gatcggtgc agaagaatg tggccgc当地 gttatcactc	2640
atggttatgg cagcactgca taattctt actgtcatgc catcgttaag atgctttct	2700
gtgactgggt agtactcaac caagtattc tgagaatagt gtatgcggcg accgagttgc	2760
tcttgccgg cgtaataacg ggataatacc gcgc当地 acata gcaactt当地 aaaagtgc当地	2820
atcattggaa aacgttcttc gggcgaaaaa ctctcaagga tcttaccgc当地 gttgagatcc	2880
agttcgatgt aacccactcg tgcacccaa acatgttcaag cgttaatatt ttgc当地	2940
gtttctgggt gagcaaaaac aggaaggcaa aatgc当地 ccaa aaaaggaa aaggcgaca	3000
cgaaaaatgtt gaataactcat actcttc当地 tttcaatatt attgaagcat ttatcagggt	3060
tattgtctca tgacggata catatttgc当地 tgatatttgc当地 aaaataaaca aatagggggtt	3120
ccgc当地 cacat ttccccggaa agtgc当地 acctt aatttgc当地 agtgc当地 atttgc当地	3180
tcgc当地 taaa ttgttgc当地 atcagctcat tttttgc当地 ataggcgaa atcggcaaaa	3240
tcccttataa atcaaaaagaa tagaccgaga taggggttgc当地 ttttgc当地 gtttggaaaca	3300
agagtccact attaaagaac gtggactcca acgtcaaaagg gc当地 aaaaacc gtctatcagg	3360
gc当地 gatggccc actacgtgaa cc当地 atc当地 aatcaagttt tttggggc当地 aggtgc当地	3420
aagcaactaaa tcggaaccct aaaggcgcc cccgatttgc当地 agttgc当地 ggaaagccgg	3480
c当地 aacgtggc gagaaaggaa gggaaaggaa cggaaaggaggc gggc当地 ctgc当地 ggc当地 ctggcaaa	3540
gtgttagcggt caccgtcgcc gtaaccacca caccggccgc gcttaatgc当地 cc当地 ctacagg	3600
gc当地 cgtccca ttccgc当地 ctttgc当地 aggctcgcc acgttgc当地 gggc当地 gatcg gtc当地 gggccct	3660
cttc当地 ctatt acgccc当地 ctgc当地 gggaaagggg gatgtgc当地 aaggcgat当地 agtgc当地 ttgc当地	3720
cgccagggtt ttcccgatca cgacgttgc当地 aaacgc当地 cggc当地 cgtat当地 ggtat当地 aactacg	3780
actactata gggc当地 atttgc当地 gagtc当地 cacc gggc当地 catgc当地 aggccacgc当地	3840
atgataagat acatttgc当地 gtttgc当地 accacaacta gaatgc当地 agtgc当地 aaaaaaaaaatgc	3900
tttatttgc当地 aaatttgc当地 tgctatttgc当地 ttatttgc当地 ccattataaag ctgc当地 aataaa	3960
caagttcttgc当地 tacagctcgcc ccatgc当地 gagtgc当地 cccgatgc当地 gggc当地 ggtca当地 cggacttccag	4020
caggaccatg tgatcgccct tctcgatggg gtc当地 ttgc当地 ctttgc当地 agggc当地 gacttgc当地 ggtatgc当地	4080
gttagtgggtt tcgggc当地 gagca gca cggggccgc gtc当地 cccgatgc当地 ggggatgc当地 tctggatgc当地	4140
gtc当地 ggccgc当地 tgacgc当地 ctgc当地 ctccgc当地 gtttgc当地 ggccatgc当地 atcttgc当地 tca ccttgc当地	4200
gc当地 ctttgc当地 tgcttgc当地 ccatgc当地 gacgttgc当地 ctgttgc当地 tagtactccag	4260
cttgc当地 gccccc aggatgttgc cgtcccttgc当地 gaagtc当地 gatgc当地 cc当地 ttgc当地 cgtatgc当地 cgggttgc当地	4320
caccagggtt tc当地 cccctc当地 acttc当地 cccctc当地 ggc当地 cggggc当地 ttgtatgc当地 cgtc当地 ctcc	4380
gaagaagatg gtgc当地 ctgc当地 ggacgttagcc ttc当地 cccgatgc当地 gccgacttgc当地 agaagtc当地	4440
ctgcttgc当地 tggc当地 cgggttgc当地 agc当地 gggc当地 gagtgc当地 cccgatgc当地 aggtggc当地	4500
gagggtgggc cagggc当地 acggc当地 gc当地 agtgc当地 ctgc当地 cccgatgc当地 ggtatgc当地	4560
gc当地 cgttaggtt gcatgc当地 cc当地 cccctc当地 ggacgc当地 ctgc当地 aacttgc当地 ggtatgc当地	4620
gc当地 cgtccagc tc当地 accaggaa tggc当地 caccac cccgatgc当地 agtgc当地 cccctc当地 ctttgc当地	4680
catggtggcg ttagcttgc当地 ttgc当地 acgtgc当地 cccgatgc当地 gggc当地 cccgatgc当地 ttgc当地	4740
agccacacgc gcaactcgcc ccataaaagg caacttgc当地 aacgc当地 cggggc当地 ggtatgc当地	4800
ttcgagtc当地 ttatttgc当地 cccctc当地 acgtgc当地 attttgc当地 cccctc当地 cccctc当地	4860
ctcccttgc当地 gggggaaaaa aaaaaaaaaa aaaaaaaaaaag gagaagagaa aaaaagcgaa	4920
cgagagagaa agcgagatgc当地 aggaagaggaa tgaagatgc当地 ggc当地 atgggttccg	4980
taggccc当地 cggacaagaa tagccccgc ctttgc当地 gagtccat local teatgtactc	5040
agatc当地 agacgc当地 tctagaacta gttgatgc当地 ataagccatgc当地 aagcactgc当地	5100
ttctctatgc当地 agccagagag ctctgc当地 ttatgc当地 atagacctcc accatgc当地 aacgc当地	5160
catttgc当地 aatggggc当地 agttgttgc当地 acatgggc当地 aagtccgc当地 gatgttgc当地	5220
ccaaaacaaa ctccattgc当地 cgtcaatggg gttgatgc当地 ggaaatcccc gttgatgc当地	5280
ccgctatcca cgc当地 cattgc当地 aataggggc当地 gtacttgc当地 tatgatac	5340
taatacgttag atgtactgc当地 aagtaggaaa gttccataag gtc当地 atgtactc当地 gggcataatgc当地	5400
ccaggcgggc catttaccgc当地 cattgc当地 aataggggc当地 gtacttgc当地 tatgatac	5460
ttgatgtact gccaatggg cagttaccgc当地 taaatactcc accatgc当地 gtcaatggaa	5520

agtccctatt	ggcgttacta	tgggaacata	cgtcattatt	gacgtcaatg	ggcgaaaaatgc	5580
gttggccggt	cagccaggcg	ggccatttac	cgtaaatgtt	gtacgcgga	actccatata	5640
tgggctatga	actaatgacc	ccgttaattga	ttactattaa	taactagtca	ataatcaatg	5700
tcaaccccg	gctgcaggaa	ttctaccggg	taggggaggc	gcttttcca	aggcagtcgt	5760
gagcatgcgc	tttagcagcc	ccgctggcac	ttggcgcac	acaagtggcc	tctggcctcg	5820
cacacattcc	acatccacccg	gtagcgccaa	ccggctccct	tctttggtgg	cccccttcgcg	5880
ccaccttcta	ctccctccct	agtcaaggaag	ttcccccccg	cccccgagct	cgcgtcgtgc	5940
aggacgtgac	aatatggaaat	agcacgttcc	actagtctcg	tgcagatgga	caagcaccgc	6000
tgagcaatgg	aaggcggttag	gcctttgggg	cagcggccaa	tagcagctt	gtcttcgc	6060
tttctgggct	cagaggctgg	gaagggggtgg	gtccggggggc	gggctcaggg	gcccggctca	6120
ggcgccggcg	ggcgcgaagg	tcctccggac	ccggcattct	gcacgcttca	aaagcgcacg	6180
tctggccgc	tgttcttc	ttcctcatct	ccgggcottt	gacctgcatg	aaaaagcctg	6240
aactcaccgc	gacgtctgtc	gagaaggttt	tgatcgaaaa	tttcgacagc	gtctccgacc	6300
tgatgcagct	ctcggagggc	gaagaatctc	gtgtttttag	tttcgatgt	ggagggcggt	6360
gatatgtcct	ggggtaat	agctgcgcgg	atgttttcta	caaagatgt	tatgtttatc	6420
ggcactttgc	atcggccgcg	ctcccgattt	cggaaatgtt	tgacatttgg	gaattcagcg	6480
agagcctgac	ctattgcata	tccgcgttgc	cacagggtgt	cacgttgc	gacctgcctg	6540
aaaccgaact	gccgcgtt	ctgcagccgg	tcgcggaggc	catggatgc	atcgctgcgg	6600
ccgatcttag	ccagacgagc	gggttcggcc	cattcggacc	gcaaggaaatc	ggtcaataca	6660
ctacatggcg	tgatccata	tgcgcgattt	ctgatcccc	tgtgtatcac	tggcaaactg	6720
tgatggacga	caccgtcagt	gcgtccgtcg	cgcaggctt	cgatgagct	atgttttggg	6780
ccgaggactg	cccccg					6795

<210> 40

<211> 5821

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic recombinant
DNA vector sequence

<400> 40

ggatctcaag	aagatcctt	gatctttct	acggggtctg	acgctcagtg	gaacgaaaac	60
tcacgttaag	ggattttgt	catgagatta	tcaaaaagga	tcttcaccta	gatccttttta	120
aattaaaaat	gaagttttaa	atcaatctaa	agtatatatg	agtaaactt	gtctgacagt	180
taccaatgt	taatcagtga	ggcacccatc	tcagcgatct	gtctatttgc	ttcatccata	240
gttgcctgac	tccccgtcgt	gtagataact	acgatacggg	agggcttacc	atctggcccc	300
agtgcgtcaa	tgataccgcg	agacccacgc	tcacccgctc	cagatttac	agcaataaaac	360
cagccagccg	gaagggccga	gcmcagaatg	ggtcctgaa	ctttatccgc	ctccatccag	420
tctattaatt	gttgcggga	agctagatgt	agtagttcgc	cagttatag	tttgcgcaac	480
gttggccca	ttgttacagg	catcggtgt	tcacgctcgt	cggttggat	ggcttcattt	540
agctccgggt	cccaacgatc	aaggcgagtt	acatgatccc	ccatgttgc	aaaaaaagcg	600
gttagctcct	tcggccctcc	gatcggtgtc	agaagtaat	tggccgcagt	gttatcactc	660
atggttatgg	cagcactgca	taattcttct	actgtcatgc	catccgtaa	atgttttct	720
gtgactgggt	agtaactcaac	caagtcatcc	tgagaatagt	gtatgcggcg	accgagttgc	780
tcttgcggcg	cgtcaatacg	ggataatacc	gcccacata	gcagaactt	aaaatgtctc	840
atcattggaa	aacgttctt	ggggcgaaaa	ctctcaagga	tcttaccgc	gttggatcc	900
agttcgatgt	aacctactcg	tgcacccaa	tgtatcttgc	catcttttac	tttcaccagg	960
gtttctgggt	gagcaaaaac	aggaaggcaa	aatgccc	aaaagggaaat	aagggcgaca	1020
cgaaatgtt	gaataactcat	actcttctt	tttcaatatt	attgaagcat	ttatcagggt	1080
tattgtctca	tgagcggata	catatttgc	tgtatttgc	aaaataaaaca	aataggggtt	1140
ccgcgcacat	ttccccggaa	agtgcaccc	gacgtcgacg	gatcgggaga	tctccgatc	1200
ccctatgggt	cactctcgt	acaatctgc	ctgatgcgc	atagtttgc	cagtatctgc	1260
tccctgcctt	tgtgttggag	gtcgctcgt	agtgcgcgag	aaaattaa	gctacaacaa	1320
ggcaaggctt	gaccgacaat	tgcatgaaga	atctgtttag	ggttaggcgt	tttgcgtcg	1380
ttcgcgtatgt	acggggccaga	tatacgcgtt	gacatttgc	attgactgt	tattaatgt	1440

aatcaattac	gggttcatta	gttcatagcc	catatatgga	gttccgcgtt	acataactta	1500
cggtaaatgg	cccgccctggc	tgaccgccc	acgacccccc	cccattgacg	tcaataatga	1560
cgtatgtcc	catagtaacg	ccaataagg	cttccattg	acgtcaatgg	gtggagtatt	1620
tacggtaaac	tgcccactt	gcagtagatc	aagtgtatca	tatgccaagt	acgcccccta	1680
ttgacgtcaa	tgacggtaaa	tggccgcct	ggcattatgc	ccagtagatc	accttatggg	1740
actttctac	ttggcagtagc	atctactat	tagtcatcgc	tattaccatg	gtgatgcgg	1800
tttggcagta	catcaatggg	cgtggatagc	ggttgactc	acggggattt	ccaagtctcc	1860
accccatg	cgtcaatggg	agttt	ggcacaaaaa	tcaacgggac	tttccaaaat	1920
gtcgtacaa	ctccgcccc	ttgacgaaa	tggccgtag	gcgtgtacgg	tggaggtct	1980
ataaagcag	agctctctgg	ctaactagag	aaccactgc	ttactggctt	atcgaattaa	2040
atacgactc	ctataggag	accaagctg	gctagcttt	aaacttaagc	ttgccaccat	2100
gaagctact	tcttctatcg	aacaagcatg	cgatattgc	cgactaaaaa	agtcgaatgt	2160
ctccaaagaa	aaaccgaagt	gcccggatg	tctgaagaac	aactgggagt	gtcgctactc	2220
tcccaaacc	aaaaggctc	cgctgactag	ggcacatctg	acagaagtgg	aatcaaggct	2280
agaaagactg	gaacagctat	ttctactat	tttcctcga	gaagacctt	acatgattt	2340
aaaatggat	tcttacagg	atataaaagc	attgttaaca	ggattattt	tacaagataa	2400
tgtgaataaa	gatggcgta	cagatagatt	ggcttcagt	gagactgata	tgccctcta	2460
attgagacag	ttgactgtat	cgccgggtac	cgagctcgga	tccactatgc	cagtgtgg	2520
gaattctgca	gatatccago	acagtggcg	ccgctcgagt	ctagaggggc	cgtttaaacc	2580
cgctgatcag	cctcgactgt	gccttctat	tgccagccat	ctgttgtt	cccccccccc	2640
gtgccttcct	tgaccctgga	aggtgcca	ccactgtcc	tttcctaata	aatgaggaa	2700
attgcatcgc	attgtctgag	taggtgtcat	tctattctgg	gggggtggg	ggggcaggac	2760
agcaaggggg	aggattggg	agacaatagc	aggcatgtg	gggatgcgg	gggctctatg	2820
gcttctgagg	cgaaaagaac	cagctgggc	tctaggggg	atccccacgc	gccctgttagc	2880
ggcgcattaa	gwgccggcgg	tgtgggtt	acgcgcagcg	tgaccgc	acttgc	2940
gccctagcgc	ccgctccctt	cgcttcttc	ccctcccttc	tcgcccacgtt	cgccggctt	3000
ccccgtcaag	ctctaaatcg	ggggctccct	ttagggttc	gatttagtgc	ttacggc	3060
ctcgacccca	aaaactatcg	ttaggggtat	ggttacgta	gtggccatc	gccctgata	3120
acggttttc	gccccttgc	gttggatc	acgttcttta	atagtggact	ctgtttccaa	3180
actggAACAA	cactcaaccc	tatctcg	tattctttt	atttataagg	gattttgc	3240
atttcgcc	attggtaaa	aatgagctg	atthaacaaa	atthaacgc	gaattaattc	3300
tgtggatgt	gtgtcagtt	gggtgtgg	agtc	ccccagg	ctccccagca	3360
tgcaaagcat	gcatctcaat	tagtcag	ccaggtgtgg	aaagtccc	ggctccccag	3420
caggcagaag	tatgcaaagc	atgcata	attagtcagc	aaccatagtc	ccgcccctaa	3480
ctccgccccat	ccggcccc	actccggcc	gttccggcc	ttctccgccc	catggctgac	3540
taattttttt	tatttatgc	gaggccgagg	ccgcctctgc	ctctgagct	ttccaga	3600
agtggagg	ctttttgg	ggcctagg	tttgc	aaaaaa	gctccggg	3660
ccattttgc	atctgatcaa	gagacaggat	gaggatc	tcgcatgatt	gaacaagatg	3720
gattgcacgc	aggttctcc	gccgcttgg	tggagagg	attcgctat	gactgggcac	3780
aacagacaat	cggtctgtct	gatggcc	tgttccgg	gtcagcgc	gggcgcccgg	3840
ttcttttgc	caagaccgac	ctgtccgg	cctg	aatg	actgcaggac	3900
ggctatcgt	gctggccacg	acgggcgtt	cttgcgc	tgtgctcgac	gttgtact	3960
aagcggaaag	ggactggctg	ctattggcg	aagtgc	gggg	gcaggatc	4020
accttgc	tgccgagaaa	gtatccat	tgc	atgcgg	ctgcata	4080
ttgatccggc	tac	tcgacc	aagc	aaaa	tcgc	4140
ctcgatgg	agccgg	tcgatc	atgat	atgg	caggg	4200
cgccagccg	acttgc	agg	tc	cg	cg	4260
tgacccatgg	cgatgc	ttg	cc	at	at	4320
tcatcgact	tggccgg	gtgtgg	acc	gtat	cg	4380
gtgatattgc	tgaagag	ggcgg	gat	gg	ttgg	4440
tcgcccgtcc	cgat	cgat	tc	tct	ctt	4500
cgggactctg	gggttgc	tgacc	agc	ac	ttt	4560
cgattccacc	ggcc	atgaa	gg	at	cc	4620
ctggatgatc	ctcc	ggat	gt	gtt	cc	4680
tattgcagct	tataatgg	aca	aa	aa	tt	4740
attttttca	ctgcatt	ttt	gg	cc	at	4800
ctgtataacc	tcgac	gt	gg	at	tc	4860
gtgaaattgt	tatcc	caat	ccaca	ca	acata	4920
					ca	

agcctggggt	gcctaattgag	ttagactaact	cacattaatt	gcgttgcgt	cactgcccgc	4980
tttcaggatcg	ggaaaacctgt	cgtgccagct	gcattaatga	atcgccaaac	gchgccccggag	5040
aggcggtttg	cgtattgggc	gctcttcgc	ttcctcgctc	actgactcgc	tgcgctcggt	5100
cgttcggtcg	cggcgagcgg	tatcagctca	ctcaaaggcg	gtaataacgg	tatccacaga	5160
atcaggggat	aacgcaggaa	agaacatgtg	agaaaaaggc	cagaaaaagg	ccaggaaccg	5220
taaaaaggcc	gctgtgctgg	cggtttcca	taggctccgc	ccccctgacg	agcatcacaa	5280
aaatcgacgc	tcaagtccaga	ggtggcgaaa	cccgacagg	ctataaagat	accaggcggt	5340
tcccccttgg	agtccttcg	tgcgcttc	tgttccgacc	ctgcccgtt	ccggatacc	5400
gtccgcctt	ctcccttcgg	gaagcgtggc	gctttctcat	agctcacgct	gtaggttatct	5460
cagtteggtg	taggtcggtt	gctccaagct	ggctgtgtg	cacgaacccc	ccgttcagcc	5520
cgaccgcgtc	gccttatccg	gtaactatcg	tcttgagtcc	aacccggtaa	gacacgactt	5580
atcgccatcg	gcagcagcca	ctggtaacag	gattagcaga	gcgaggatag	taggcgggtgc	5640
tacagagttc	ttgaagtgg	ggcctaacta	cgcgtacact	agaagaacag	tatttggtat	5700
ctgcgctctg	ctgaagccag	ttaccttcgg	aaaaagagtt	ggtagctt	gatccggcaa	5760
acaaccacc	gctggtagcg	gttttttgt	ttgcaagcag	cagattacgc	gcagaaaaaa	5820
a						5821

<210> 41
<211> 10075
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant
DNA molecule sequence

<220>
<221> modified_base
<222> (9040)..(9040)
<223> unsequenced nucleotide

<400> 41						
ctggtagcg	tggttttttt	gtttcaagc	agcagattac	gcccggaaaa	aaaggatctc	60
aagaagatcc	tttgatcttt	tctacggggt	ctgacgctca	gtggAACGAA	aactcacgtt	120
aagggatttt	ggtcatgaga	ttatcaaaaa	ggatcttcac	ctagatcctt	ttaaattaaa	180
aatgaagttt	taaatcaato	taaagtatat	atgagtaaac	ttggctgtac	agttaccaat	240
gcttaatcag	tgaggcacct	atctcagcga	tctgtctatt	tcgttcatcc	atagttgcct	300
gactccccgt	cgtgttagata	actacgatac	gggagggctt	accatctggc	cccagtgcgt	360
caatgatacc	gcgagaccca	cgctcaccgg	ctccagattt	atcagcaata	aaccagccag	420
ccggaagggc	cgagcgcaga	agtggctctg	caacttttac	cgccctccatc	cagtcttata	480
attgttgcgg	ggaagctaga	gtaagttagt	cgccagttaa	tagttgcgc	aacgttggtt	540
ccattgtac	aggcatcg	gtgtcacgt	cgctcggttg	tatggcttca	ttcagctccg	600
gttcccaacg	atcaaggcga	gttacatgt	cccccatgtt	gtcaaaaaaa	ggggttagct	660
ccttcggtcc	tccgatcg	gtcagaagta	agttggccgc	agtgttatca	ctcatggta	720
tggcagcact	gcataattct	cttactgtca	tgcacatccgt	aagatgtttt	tctgtgactg	780
gtgagtactc	aaccaagtca	ttctgagaat	agtgtatgcg	gcccggagg	tgctcttgcc	840
cggcgtcaat	acgggataat	accgcggccac	atagcagaac	tttaaaatgt	ctcatcattt	900
aaaaacgttc	tccggggcga	aaactctaa	ggatcttacc	gctgttggaa	tccagttcga	960
tgttaaccac	tcgtgcaccc	aactgtt	cagcatctt	tacttttacc	agcgtttctg	1020
ggtgagcaaa	aacaggaagg	caaaatcccg	caaaaaagg	aataaggcg	acacggaaat	1080
gttgaatact	catactctt	cttttcaat	attattgaag	catttatcg	ggttattgtc	1140
tcatgagcgg	atacatattt	gaatgtattt	agaaaaataa	acaaataggg	gttccgcgc	1200
catttcccc	aaaagtgc	cctaaattgt	aagcgtt	atttgtt	aattcgcgtt	1260
aaattttgt	taaatcagct	catttttaa	ccataggcc	gaaatccgg	aaatccctt	1320
taaatcaaaa	gaatagaccg	agatagggtt	gagtgtt	ccagttgg	acaagagtcc	1380
actattaaag	aacgtggact	ccaaacgtca	agggcgaaaa	accgtctatc	agggcgatgg	1440
cccactacgt	gaaccatcac	cctaattcaag	tttttgggg	tcgagg	tgaaagcact	1500

aatcggaaac	cctaaaggga	gcccccgatt	tagagcttga	cggggaaaagc	ccgcgaacgt	1560
ggcgagaaaag	gaagggaaga	aagcgaaaagg	agcggggcgct	agggcgctgg	caagtgtagc	1620
ggtcacgctg	cgcgtaaacca	ccacaccgc	cgcgttaat	gcccgcctac	agggcgcgctc	1680
ccattcgcca	ttagggctgc	gcaactgttg	ggaaggcga	tcggtgccgg	cctcttcgct	1740
attacgcag	ctggcgaaag	ggggatgtgc	tgcaggcga	ttaagttggg	taacgccagg	1800
gttttcccgag	tcacgacgtt	gtaaaacgac	ggccagttag	cgcgcgtaat	acgactcact	1860
atagggcgaa	ttggagctcc	accgcgtgc	ggccgggcca	tgcaggccac	gacatgataa	1920
gatacattga	tgagtttgga	caaaccacaa	ctagaatgca	tgaaaaaaa	tgctttattt	1980
gtgaaatttg	tgtgttattt	gtcttatttg	taaccattat	aagctgcaat	aaacaagttc	2040
ttgtacagct	cgtccatgcc	gagagtgtatc	ccggcgccgg	tcacgaactc	cagcaggacc	2100
atgtgatcgc	gcttctcggtt	ggggcttttg	ctcaggcgg	actggtagct	caggttagtgg	2160
ttgtcgccgca	gcagcacggg	gccgtcgccg	atgggggtgt	tctgttgta	gtggtcggcg	2220
agctgcacgc	tgccgtcctc	gatgttgtgg	cgcgttgcga	agttcacctt	gatgccgttc	2280
ttctgtctgt	cggccatgtat	atagacgttg	tggtgttgc	agttgtactc	cagctgtgc	2340
cccaggatgt	tgcgtcctc	cttgaagtgc	atgccttca	gctcgatgcg	gttcaccagg	2400
gtgtcgccct	cgaacttcac	ctcggcgccg	gtcttgcgt	tgccgtcgtc	cttgaagaag	2460
atggtgccgt	cctggacgta	gccttcggcc	atggggact	tgaagaagtc	tggtgtcttc	2520
atgtgttcgg	ggtagcgccg	gaagcactgc	aggccgtac	cgaagggtgt	cacgagggtg	2580
ggccaggggca	cggcgacgtt	gccgggtgg	cagatgaact	tcagggtcag	tttgcgttag	2640
gtggcattcgc	cctcgccctc	gccggacacg	ctgaacttgc	ggccgttac	gtcggcggtcc	2700
agctcgacca	ggatgggcac	caccccggt	aacagcttc	cgcccttgc	caccatgggt	2760
gcgttagctt	gatttgacag	tggctgggg	ttgcgcgc	gggtttata	ggaagccaca	2820
gcggccactc	gagccataaaa	aggcaactt	agaacggcg	gggggtgatt	ggattcgagt	2880
cgtttattca	ccggccttgc	cgcacagtgc	agcattttt	tacccctct	ccctctt	2940
tgccggggaa	aaaaaaaaaa	aaaaaaaaaa	aaggagaaga	aaaaaaaaagc	gagcgagaga	3000
gaaagcgaga	tttaggaaga	ggatgaagag	tttgcgtat	gggtctgggt	ccgtaggccc	3060
agatggacaa	gaatagcccc	cgccttgcg	gacagtatcc	cattcagtga	ctcagatca	3120
atcaagcgcc	cgcgtgttgc	atggatatct	gcagaattcc	agcacactgg	cgccgttac	3180
tagtggatcc	ctctggagct	gcctagtgt	ggtgtctgc	ggaaacttct	aaaatgtctc	3240
ggcctattta	atgggagaat	tcactgaaga	attaaaggac	tcagtaaaga	agaggtttgg	3300
tttggatcatg	aaagtaagta	gcttgcgt	tgctttattt	tcctgtgact	aaaattccac	3360
cttgcatttc	ccaaataagga	ttgcttcaag	ctatthaata	caaattaaac	ctgtttcacc	3420
atgctcatag	gaaaatgtca	aagtatctc	attcacaac	tgagggtct	ctttcttt	3480
aaggtagctc	tgagatttgc	tttgcataatgg	agacaaagac	ttatttaact	gttggggagg	3540
gggaaggagg	acttaggc	ggcaggcg	gatgtgggg	acgaaccatg	aggggcaatt	3600
tggctgcctc	tgaaaagggtc	cagagctacc	ttaaaagaca	ggcttctat	aacaagtgtt	3660
ccttctttt	ccttctttt	ttttttgtag	acagagtctc	actctgtgc	ccaagctgga	3720
gtgcagttgt	gcgtatctcg	ctcaactgca	cctccagctc	ccaggttcaa	gcaattctcc	3780
tgcctcagcc	tcccaagtag	ctgggattac	aggcaaccgc	taccatgccc	agctaatttt	3840
ttgtatTTT	aatagagatg	gggttccacc	atgttgc	gggtgttctc	aaactcctga	3900
cctcaagtga	tccacccacc	tcagccccc	aaagtgtcg	gattacagac	atgagccacc	3960
gcgcccacac	acagctggct	ttgtcgccca	ggctggaagt	gcagttggc	gatctcggt	4020
cactgcaagc	tctgcctcct	gggttcaacgc	cattctctcg	cctcagcttc	ccaaagttagct	4080
gagactacag	gcccggccca	ccacgccccg	ctaattttt	gtatTTT	tgtagatgtt	4140
gtttcaccat	gttagccagg	atggcttgc	tctctgcacc	tcgtgatccg	cctgttctgg	4200
cctcccaaaag	tgtgggttgc	acaggctga	gcccacacgc	ctggcttaca	agtggccctt	4260
tctaaggcccc	aagatgtcac	tgtctttca	ctgaaggcg	aaaaagaatg	taggcatttt	4320
tgtctgagag	ggaggggtat	gttttaggag	aaaggacca	gacacattt	tgagttggaa	4380
gtaagtgggtt	tcctttctca	gccttatgtat	atggtttgc	tctgttccc	cacccaaatc	4440
tcatggtcaa	ctgtatcccc	cagtggtga	ggtagggct	ggcggggagg	actggatcat	4500
gggggtgtatc	cttcatgtat	ggtttagcac	cattccttgc	gtgtactct	agtgtatcg	4560
ttctcacaag	atctgtatgg	gtaaaagtgt	gtgtcttgc	ttttgtctcc	tgccatgc	4620
gacgcctact	cccaacttcac	cttctgtat	gagtcaaatt	ccctgaggcc	ttccccagaa	4680
cagaagcttc	ctgtacagcc	ttgtggccca	cgagccaaatt	aaaccttctt	tctttataaa	4740
ttaccaggc	tcagggattt	ctttatagca	gactaaataca	cctcaagta	gggattttatt	4800
tttattttct	attttttttt	tgagataggg	tctcgatctg	tcacccaggc	ttggagtgcag	4860
tggcatgatc	tcagcttgc	gtaaacctctg	cctccgggtt	tgaagcaatt	ctcatgcct	4920
aqccacccaa	qtagctggaa	ttacagggt	gtggccaccc	gctcagctaa	ttttgtat	4980

ttagtagag	ttgggttgc	tccatgttg	ccaggctgg	atcaaactcc	tggactcaag	5040
caatctgccc	accttggcct	ccccagagt	ccaggattac	aggcatgacc	catcttcgcct	5100
ggtcagttt	tttttttta	aatccagtt	atgctgggt	ccatgactca	cacctgtaat	5160
gccagtgcct	tgggaggccg	agggtggaaag	atcaacttgc	gccaggagtt	tgagaccagc	5220
ctggaaaca	tctcatctc	acaaaaaac	aaaaacaaaa	aaaacaaca	acaacatcaa	5280
acataggccag	gtgtgggt	gcacacctgt	agtcccagct	ccttggggagg	ctgaggcgagg	5340
aggattgttt	gagccagga	gttcaaggct	gcagtgaact	atgattgcac	cactgcattc	5400
cagcctggc	aacagagcaa	gaacctatct	ctattaaga	aaaatgtct	ctaagaaaatc	5460
atcttagtac	tttgtaact	attaagttt	tttttaagt	tattaattac	atagccagta	5520
taataactta	gttgcataa	gctttgagg	ggttccaagc	aagaagaaaag	gagttgttgt	5580
taagctacag	ctcagactct	cagaatgaac	tgttccatc	ttcgctgtcc	tttgcttaac	5640
tgtggataa	ataataattt	ctcaacatcc	aacacatgt	tccttaatcg	tcctctacga	5700
cactgggtcc	atgtgccaa	gagaacccag	aacagccgca	aacacctact	ggattgttac	5760
tggccttcc	tgagaggat	acccggcctt	tgaccttgc	catgttgaag	tttttcaagt	5820
aggcctcgta	gattctctt	gccagagatt	tgagatctgc	agtttcagaa	tcttctatgt	5880
catgttcaca	ggtaagaatt	tctgtttca	gtttgttctt	ctcagatctt	ggcattcgtc	5940
cggtaaccgg	cgatacagtc	aactgtcttt	gaccttgc	actactctct	tccgatgtat	6000
atgtcgact	tattctatgc	tgtctcaatg	tttagaggcat	atcagtcctc	actgaagcca	6060
atctatctgt	gacggcatct	ttattcacat	tatcttgcac	aaataatcct	gttaacaatg	6120
cttttatatc	ctgttaagaa	tccatttca	aaatcatgtc	aaggtcttct	cgaggaaaaaa	6180
tcagtagaaa	tagctgttcc	agtctttca	gccttgattc	cacttctgc	agatgtgcct	6240
tagtcagcgg	agacacccat	gttttggag	agtagcgaca	ctcccaagtt	tttttcaagac	6300
acttggcgca	cttcgggtt	tctttggagc	acttgagctt	ttaagtgcgg	caaataatcg	6360
atgcttgc	gatagaagac	agtagcttca	tggtggcaag	cttaagttt	aacgctagaa	6420
ctagtgatc	cggtataagcc	agtaagcagt	gggttctct	gttagccaga	gagctctgt	6480
tatatagacc	tccaccatgt	cacgcctacc	gcccatattgc	gtcaatgggg	cgaggatgtt	6540
acgacat	ggaaagtccc	gttggat	gtggccaaac	aaactccat	tgacgtcaat	6600
gggggtggaga	cttggaaatc	ccctgtggc	aaaccgttat	ccacgccccat	tgatgtactg	6660
ccaaaaccgc	atcaccatgg	taataggat	gactaatac	tagatgtact	gccaagtagg	6720
aaagtccat	aaaggcatgt	actggcata	atgcaggcg	ggccatttac	cgtcattgac	6780
gtcaataggg	ggcgctactt	gcatatgata	cacttgatgt	actgccaagt	gggcagttt	6840
ccgtaaatac	tccacccatt	gacgtcaatg	gaaagtccct	attggcgtt	ctatggaaac	6900
atacgtcatt	attgtacgtca	atggccggg	gtcggtggc	gttcagccag	gccccccatt	6960
tacgttaagt	tatgtaacgc	ggaactccat	atatggcata	tgaactaatt	accccgtaat	7020
tgattactat	taataactag	tcaataatca	atgtcaaccc	cgggctgc	gaattctacc	7080
gggttagggg	ggcgctttc	ccaaggcagt	ctggagcat	cgctttagca	gccccgctgg	7140
cacttggcgc	atcacaatgt	gcctctggcc	tcgcacacat	tccacatcca	ccggtagcgc	7200
caaccggctc	ccttcttgg	tggcccttc	gcgcacatt	ctactcctcc	cctagtcagg	7260
aagttccccc	ccgccccgca	gtcgcgtcg	tgaggacgt	gacaaatgg	agtagcacgt	7320
ctcaactagtc	tcgtgcagat	ggacaagcac	cgctgagcaa	tggaaagcgg	taggccttgc	7380
gggcagcggc	caatagcgc	ttggcttcc	cgcttctgg	gctcagaggc	tggaaagggg	7440
tgggtccggg	ggcggtctca	ggggcggt	cagggcgccc	gcggcgcg	aggtcctccg	7500
gaccggcat	tctgcacgt	tcaaaagcgc	acgtctgc	cgctgttctc	ctcttcctca	7560
tctccgggccc	ttcgacactgc	atgaaaaagc	ctgaactc	cgcgacgtct	gtcgagaagt	7620
ttctgtatcg	aaagttcgac	agcgtctcc	acctgatgc	gctctcgag	ggcgaagaat	7680
ctcggtctt	cagcttcgtat	gtaggaggc	gtggatatgt	cctgogggt	aatagctgcg	7740
ccgatggttt	ctacaaagat	cgttatgtt	atcggactt	tgcattcg	cgctcccgg	7800
ttccggaaat	gcttgacatt	ggggaaatca	gagagacgt	gacatttgc	atctcccgcc	7860
gtgcacagg	tgtcacgtt	caagacatgc	ctgaaaccga	actgccccgt	gttctgc	7920
cggtcgggaa	ggccatggat	cgcatcgct	cgccgatct	tagccagac	agcgggttc	7980
gcccatcg	accgcaagga	atcggtcaat	acactacat	gcgtgattt	atatgcgcg	8040
ttgctgtatcc	ccatgtgtat	cactggaaa	ctgtgtat	cgacaccgtc	agtgcgtcc	8100
tgcgcgaggc	tctcgatgag	ctgatgttt	ggggcgagga	ctgccccgaa	gtccggcacc	8160
tcgtgcacgc	ggatttcggc	tccaacaatg	tcctgacgg	caatggccgc	ataaacagcg	8220
tcattgtact	gagcgaggcg	atgttccggg	atttccata	cgaggtcgcc	aacatcttct	8280
tctggaggcc	gtgggtggct	tgtatggac	agcagacgc	ctacttcgag	cgaggcgatc	8340
cggagcttgc	aggatcgccg	cggctccggg	cgtatatgt	ccgcatttgg	cttgaccaac	8400
tctatcagag	cttgggttgc	ggcaatttgc	atgtatgc	cgac	ttggcgccag	8460

acgcaatcgt	ccgatccgga	gccgggactg	tcgggcgtac	acaaatcgcc	cgcagaagcg	8520
cggccgtctg	gaccgatggc	tgtgtagaag	tactcgccga	tagtggaaac	cgacgccccca	8580
gcactcgcc	gagggcaaag	gaatgctga	gaaaggaagt	gagctgtaaa	ggctgagctc	8640
tctctctgac	gtatgttagc	tctggtagc	ttcgtcaactc	actgttcttg	actcagcatg	8700
gcaatctgtat	gaaatcccag	ctgttaagtct	gcataaattt	atgatctatt	aaacaataaa	8760
gatgtccact	aaaatggaaag	ttttttactg	tcatactttt	ttaagaaggg	tgagaacaga	8820
gtacctacat	tttgaatgg	aggattggag	ctacgggggt	gggggtgggg	gtgggattag	8880
ataaatgcct	gctcttact	gaaggcttctt	tactattgt	ttatgataat	gttctatagt	8940
tggatatcat	aattttaaaca	agaaaaacca	aattaaggc	cagctcattc	ctccactcat	9000
gatctataga	tctatagatc	tctcgtggga	tcattgttt	tctgatccac	tggaaagctt	9060
tcgataccgt	cgacctcgag	ggggggcccg	gtacccagct	tttgttccct	ttagtgaggg	9120
ttaattgcgc	gcttggcgta	atcatggta	tagctgtttc	ctgtgtaaa	ttgttatccg	9180
ctcacaattc	cacacaacat	acgagccgga	agcataaagt	gtaaagcctg	gggtgcctaa	9240
tgagtgagct	aactcacatt	aattgcgtt	cgcctactgc	ccgccttcca	gtcgggaaac	9300
ctgtcgtgcc	agctgcatta	atgaatccgc	caacgcgcgg	ggagaggggg	tttgcgtatt	9360
ggcgctt	ccgcttcctc	gctcaactgac	tcgctgcgt	cggtcggtcg	gtgcggcgca	9420
gcggtatcag	ctcaactcaa	ggcggtataa	cggttatcca	cagaatcagg	ggataaacgca	9480
ggaaagaaca	tgtgagcaa	aggccagcaa	aaggccagga	accgtaaaaa	ggccgcgttg	9540
ctggcgcccc	tccataggct	ccgcccccc	gacgagcatc	acaaaaatcg	acgctcaagt	9600
cagaggtggc	gaaacccgac	aggactataa	agataccagg	cgtttcccc	tggaaagctcc	9660
ctcggtcgct	ctcctgttcc	gaccctggcg	cttaccggat	acctgtccgc	ctttctccct	9720
tcgggaagcg	tggcgcttcc	tcatagctca	cgcgttaggt	atctcagttc	ggtgttaggtc	9780
gttcgctcca	agctgggctg	tgtgcacgaa	ccccccgttc	agcccgaccg	ctgcgcctta	9840
tccggtaact	atcgcttga	gtccaaaccc	gtaagacacg	acttacgc	actggcagca	9900
gccactggta	acaggattag	cagagccagg	tatgttaggc	gtgctacaga	gttcttgaag	9960
tggtggccta	actacggcta	cactagaagg	acagtattt	gtatctgc	tctgctgaag	10020
ccagttacct	tcggaaaaaaag	agttggtagc	tcttgcaccc	gcaaaacaaac	caccc	10075

<210> 42
<211> 8971
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant DNA vector sequence

<400> 42						
gtaggtgtca	ttctattctg	gggggtgggg	tggggcagga	cagcaagggg	gaggattggg	60
aagacaatag	caggcatgct	ggggatgcgg	tggctctat	ggcttctgag	goggaaagaa	120
ccagctgggg	ctctaggggg	tatccccacg	cgcctgttag	cgccgcattt	agcgcggcgg	180
gtgtgggtgt	tacgcgcagc	gtgaccgcta	cacttgcag	cgccctagcg	cccgctccct	240
tcgctttttt	cccttcctt	ctcgccacgt	tcgcccgtt	tccccgtcaa	gtctaaatc	300
gggggctccc	tttagggttc	cgatttagt	cttacggca	cctcgacccc	aaaaaacttg	360
attagggtga	tggttcacgt	agtggccat	cgcctgtata	gacggttttt	cgccttttga	420
cgttggagtc	cacgttctt	aatagtggac	tcttgttcca	aactggaaaca	acactcaacc	480
ctatctcggt	ctattttttt	gatttataag	ggattttgcc	gatttccggc	tattggttaa	540
aaaatgagct	gatttaacaa	aaatttaacg	cgaatttaatt	ctgtggatg	tgtgtcgtt	600
agggtgttga	aagtccccag	gctccccagc	aggcagaagt	atgcaaaagca	tgcattctaa	660
ttagtcagca	accaggtgt	gaaagtcccc	agctccccc	gcaggcagaa	gtatgaaag	720
catgcatttc	aatttagtca	caaccatagt	ccgccttcca	actccgc	tccgccttcca	780
aactccgccc	agttccgccc	attctccgccc	ccatggctga	ctaattttt	ttatttatgc	840
agaggccgag	gccgcctctg	cctctgagct	atccagaag	tagtggagg	gttttttgg	900
aggcctaggc	ttttgcaaaa	agctccggg	agcttgtata	tccattttcg	gatctgatca	960
gcacgtgatg	aaaaagcctg	aactcaccgc	gacgtctgtc	gagaagttt	tgatcgaaaa	1020
gttcgacagc	gtctccgacc	tgtatgcagct	ctcgaggggc	gaagaatctc	gtgcttca	1080
cttcgatgt	ggagggcggt	gatatgtcct	gccccgtaaat	agctgcgcgg	atggtttcta	1140

caaagatcg	tatgtttatc	ggcacttgc	atcgccgcg	ctcccgattc	cggaaagtgc	1200
tgacattgg	gaattcagcg	agagcctgac	ctattgcata	tcccgcgtg	cacagggtgt	1260
cacgttcaa	gacctgcctg	aaaccgaact	gcccgctgtt	ctgcagccgg	tgcggaggc	1320
catggatcg	atcgctgcgg	ccgatcttag	ccagacgagc	gggttcggcc	cattcggacc	1380
gcaaggaaatc	ggtcaataca	ctacatggcg	tgatttcata	tgcgcgattt	ctgatcccc	1440
tgtgtatcac	tggcaaactg	tgatggacga	caccgtcagt	gcgtccgtcg	cgcaggctct	1500
cgatgagctg	atgtttggg	ccgaggactg	ccccgaagt	cggcacctcg	tgcacgcgg	1560
tttcggctcc	aacaatgtcc	tgacggacaa	tgccgcata	acagcggtca	ttgactggag	1620
cgaggcgtatc	ttcggggatt	cccaatacga	ggtgcggaaac	atcttcttct	ggaggccgtg	1680
gttggcttgt	atggagcagc	agacgcgtca	cttcgagcg	aggcatccgg	agcttgcagg	1740
atcgccgcgg	ctccggcggt	atatgcctcg	cattggcttt	gaccaactct	atcagagctt	1800
ggttgcacggc	aatttcgtatc	atgcagcttg	ggcgcagggt	cgatgcgacg	aatcgtccg	1860
atccggagcc	gggactgtcg	ggcgtacaca	aatcgcccgc	agaagcgcgg	ccgtctggac	1920
cgatggctgt	gtagaagttac	tcgcccata	tgaaaaccga	cgccccagca	ctcgccgag	1980
ggcaaaggaa	tagcacgtgc	tacgagattt	cgattccacc	gccgccttct	atgaaagggtt	2040
gggcttcgga	atcgtttcc	gggacgcgg	ctggatgatc	ctccagcgcg	ggatctcat	2100
gctggagttc	ttcgcccacc	ccaactgtt	tattgcagct	tataatggtt	acaaataaaag	2160
caatagacatc	acaaatttca	caaataaaagc	attttttca	ctgcattcta	gttgggttt	2220
gtccaaactc	atcaatgtat	cttatcatgt	ctgtataccg	tcgacccctca	getagagctt	2280
ggcgtaatca	tggtcatagc	tgtttcctgt	gtgaaattgt	tatccgctca	caattccaca	2340
caacatacga	gccggaagca	taaagtgtaa	agcctgggt	gcctaattag	tgagctaact	2400
cacattaatt	gcgttgcgt	caactgcgc	tttccagtcg	ggaaacctgt	cgtgccagct	2460
gcattaatga	atcgcccaac	gchgccccag	aggcggttt	cgtattggc	gtcttccgc	2520
ttcctcgctc	actgactcgc	tgcgctcggt	cgttcggctg	cgggcgacgg	tatcagctca	2580
ctcaaaggcg	gtatacaggt	tatccacaga	atcaggggat	aacgcaggaa	agaacatgtg	2640
agcaaaaggc	cagcaaaagg	ccaggaacgg	taaaaaggcc	gcgttgcgtgg	cgtttttcca	2700
taggctccgc	ccccctgacg	agcatcacaa	aaatcgacgc	tcaagtca	gttggcgaaa	2760
cccgacagga	ctataaagat	accaggcggt	tccccctgg	agctccctcg	tgcgtctcc	2820
tgttccgacc	ctgcccgtta	ccggataacct	gttccctt	cttcccttcgg	gaagcgtggc	2880
gctttctcat	agtcacgct	gttaggtatct	cagttcgggt	taggtcgttc	gtccaaagct	2940
gggctgtgt	cacgaacccc	cggttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	3000
tcttgagtcc	aaccggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	3060
gattagcaga	gcgaggtatg	taggcgtgc	tacagagttc	ttgaagtgg	ggccctaacta	3120
cggttacact	agaagaacag	tatttgtat	ctgcgtctg	ctgaagccag	ttaccttcg	3180
aaaaagagtt	ggtagctctt	gatccgca	acaaaccacc	gctggtagcg	gtttttttgt	3240
ttgcaagcag	cagattacgc	cgaaaaaaa	aggatctaa	gaagatcctt	tgatctttc	3300
tacgggtct	gacgctcgt	ggaacgaaaa	ctcacgtta	gggattttgg	tcatgagatt	3360
atcaaaaagg	atcttccac	agatcccttt	aaattaaaaa	tgaagttta	aatcaatcta	3420
aagtatatat	gagtaaactt	ggtctgacag	ttaccaatgc	ttaatcagt	aggcacctat	3480
ctcagcgatc	tgtctatttc	gttcatccat	attgcctga	ctcccccgtcg	ttagataaac	3540
tacgatacgg	gagggcttac	catctggccc	cagtgcgtca	atgataccgc	gagacccacg	3600
ctcaccggct	ccagatttat	cagcaataaa	ccagccagcc	ggaaggcccg	agcgcagaaag	3660
ttgtccctgca	actttatccg	cctccatcca	gtctattaaat	tgttgcgggg	aaactagagt	3720
aagttagttcg	ccagttata	gtttgcgca	cgttgcgtcc	attgctacag	gatcgtgg	3780
gtcacgctcg	tgtttggta	tggtttcatt	cagctccgg	tcccaacat	caaggcgagt	3840
tacatgatcc	cccatgttgc	gcaaaaaaggc	ggttagctcc	ttcggctctc	cgatcggtt	3900
cagaagtaag	ttggccgcag	tgttatact	catggttatg	gcagcactgc	ataattctct	3960
tactgtcatg	ccatccgtaa	gatgctttc	tgtgactgtt	gagtaactaa	ccaaagtctt	4020
ctgagaatag	tgtatgcggc	gaccgagtt	ctttgcggcc	cggtcaatac	gggataatac	4080
cgcgcacat	agcagaactt	taaaaggtct	catcattgg	aaacgttctt	cgggcgaaa	4140
actctcaagg	atcttaccgc	tgttgcgtatc	cagttcgtat	taacccactc	gtgcacccaa	4200
ctgatcttca	gcatcttttca	ctttcaccag	cgttctgggg	tgagaaaaaa	caggaaggca	4260
aaatgccca	aaaaaggggaa	taaggcgcac	acggaaatgt	tgaataactca	tactcttcct	4320
ttttcaatat	tattgaagca	tttatcaggg	ttattgtctc	atgagcggat	acatatttga	4380
atgttatttag	aaaaataaaac	aaatagggg	tccgcgcaca	tttccccgaa	aagtgcaccc	4440
tgacgtcgac	ggatcggggag	atctcccgat	cccttatgg	gcactctcg	tacaatctgc	4500
tctgatgccc	catagttaa	ccagttatcg	ctccctgtt	gtgtgttgg	ggtcgtcgag	4560
tagtgcgcga	gcaaaaattta	agctacaaca	aggcaaggct	tgaccgacaa	ttgcgtcgaa	4620

aatctgccta	gggttaggcg	tttgcgctg	cttcgcgatg	tacgggccag	atatacgcgt	4680
tgacatttat	tattgactag	ttattaatag	taatcaatta	cggggtcatt	agttcatagc	4740
ccatatatgg	agttccgcgt	tacataactt	acggtaaatg	gccgcctgg	ctgaccgccc	4800
aacgacccccc	gcccattgac	gtcaataatg	acgtatgttc	ccatagtaac	gccaataggg	4860
actttccatt	gacgtcaatg	ggtggagtat	ttaccgtaaa	ctgcccactt	ggcagttacat	4920
caagtgtatc	atatgccaag	tacgccccct	attgacgtca	atgacggtaa	atggcccgcc	4980
tggcattatg	cccagtagat	gaccttatgg	gactttctca	cttggcagta	catctacgta	5040
ttagtcatcg	cttattaccat	ggtgatgcgg	tttggcagtt	acatcaatgg	ggtggatag	5100
cggtttgcgt	cacggggatt	tccaagtctc	caccccatgg	acgtcaatgg	gagtttgttt	5160
tggcacaaaa	atcaacggga	cttccaaaaa	tgtcgtaaca	actccgcccc	attgacgcaa	5220
atgggcggta	ggcgtgtacg	gtgggaggtc	tatataagca	gagctctctg	gctaactaga	5280
gaaccactg	cttactggct	tatcgaaatt	aatacgactc	actataggga	gacccaagct	5340
ggcttagcgtt	taaacttaag	cttgccacca	tgaagctact	gtcttctatc	gaacaagcat	5400
gcgatatttgc	ccgacttaaa	aagctcaagt	gctccaaaga	aaaaccgaag	tgcgccaagt	5460
gtctgaagaa	caactggag	tgtcgctact	ctcccaaaaac	caaaggctt	ccgctgacta	5520
gggcacatct	gacagaagtg	gaatcaagge	tagaaagact	ggaacagctt	tttctactga	5580
tttttcctcg	agaagacctt	gacatgattt	tgaaaatgg	ttctttacag	gatataaaaag	5640
cattgttaac	aggattattt	gtacaagata	atgtgaataa	agatggcgtc	acagatagat	5700
tggcttcagt	ggagactgtat	atgcctctaa	cattgagaca	gcatagaata	agtgcgacat	5760
catcatcgga	agagagtagt	aacaaaggtc	aaagacagtt	gactgtatcg	ccgggttaccc	5820
cagtgggat	gtctcataat	ggtaagtaaa	cagtcatcac	cataactttt	attattctca	5880
ttatagctgc	cagaccagtg	gacactaaag	ccattgccaa	aatgtgtac	agtttttcca	5940
ccaaatgcca	gaatttagaa	tattgcatgg	cgataaaaaca	tttctctttt	aggtcagtgt	6000
ttttaaagtt	ttattataga	acctttctct	ctgtggttgg	gcatctgcca	tgaggagaaa	6060
agagacttga	aaaatctggg	ggattatggg	aaaaacttta	gtgttagaac	aatataatgg	6120
atcatcatga	ttaattttta	tattgcctca	tgcattaaaa	acagaattaa	aattttacttt	6180
tgtgtatcag	tagtgcatt	ctgctttaaa	tgtaaagaaaa	tcctgaagat	gtctgggtct	6240
gaaactttaa	atggcatatg	agaattttaga	ttcattttat	tttttttttt	tttttttttt	6300
tgctgttat	tatgcaaaa	cgagtttcg	tttccttcac	ttcaaggaaac	gtccatgagt	6360
tgtggaaagg	gtaaaaacaaa	tttattttat	agttctgcac	aacttcaaag	ccatagaaac	6420
ataagagaat	aaaactttct	acttttggaa	cctgtgtcta	aagtcaagct	tatgacttttgc	6480
ggtgtcaact	taagggaaat	gatgtagaag	acctgagaaaa	agaccatgt	tttcatgatc	6540
tgtggacta	actcagtaag	gtcctggaag	acaaaacgtgt	ttgaaaaaat	aatttttgg	6600
cactatttgg	ttctgattgg	ttgcctggct	gacaaaataaa	attttaggtt	taacttgata	6660
gtattatatg	tgtctagttat	taatcagttt	tttcaatgc	actcaaggaa	taaaaaatgc	6720
caacatgatc	aaattaaattt	tgaagttca	tttcatcagc	tatccttattt	tttctttctt	6780
ttgttttcat	ctctctcaaa	actgatctag	taagaaggca	gtcttcatct	atatgtactc	6840
taaagcatct	gagaaaattt	aatttttgc	ctatcatgca	gatacttgg	tttcacagtg	6900
gaaaccatca	agatatttga	gagtcactca	ttgattaata	acctaccctt	taaaaaatagt	6960
tcacaaaattt	ctatctgtat	gattactaat	ggtcagttc	ctctaaaggc	aagaatgtat	7020
tttgaataga	gctatcagca	gtactgaaca	ggatggaaat	ttggtccaac	tttcctcaac	7080
aattgactac	atctgctgtat	tacaactgta	gtcatggatt	gcaaaaataat	ttgtctggttc	7140
agttccatca	gtttccctttt	tatttcagac	ctcataaaac	ataaggcga	gcttccaagt	7200
caagggttgt	attttaggtt	atgcaatcaa	gttgcataaa	acttatttca	tcttcatgaa	7260
gaaatcatttgc	acttttaggtt	cattatcact	gagacagcct	aggaaagcat	gccagttccg	7320
cttttaacac	taatttttgc	ctattaaaaa	attatcactt	cacacattca	tgtaaaaatat	7380
aattccctgtt	acttacaggc	agtgcataca	ttgatgtaaa	atatatttcc	ttgtgttttac	7440
aggcatttgc	tacgtttact	taaaatcaat	ctgcaatttgc	ttattgtttt	taggcaaaata	7500
ttatattaaac	cccaacatctt	gtatgttgg	tgttataattt	tttccttcccc	ataaaatgtca	7560
aagaaagcta	tcataaactt	gataggctat	attttgactt	catacgaaaa	tttattttat	7620
ttcttagagat	ggtcaatctt	gatagggagt	gtttgttcag	ttctagcaaa	agcaagaaaa	7680
tatataatgtat	aagtgtcatg	acatcagaga	cttaggcagt	tcaacactta	aaaaatttgc	7740
aacataacta	agaaaaccttgc	agagttttt	atgggagttt	gtttgtcga	cagaaaataaa	7800
acatcagact	aacaaacaagg	ttttgatttag	tttgcctgt	tagctgtatg	aagaaaacat	7860
ttctaaatac	aaagttttct	aggaaggta	gtcacttattt	aaaatccatt	gttttgacca	7920
atgtgaagcc	aacagggtttt	tcattcatat	agaatgcctt	tttataaaaa	gaggctctga	7980
gtgttataacc	aaaatcagca	cccagctgtt	ttcagttatgc	gcaagaaaac	atgaacgtcc	8040
cccaaagcaa	atgtactcca	ttccgtgtta	accttatttc	tgctccctgg	tctctgatta	8100

gtctggttga	tgaatgttagc	ttcccccttg	gccatgcact	cacagcttg	aaggaggctt	8160
ctgaaggaac	tggtagggaa	cccttagggc	tcttgcataat	gttgttggat	agtcaactgaa	8220
atcagtgtatc	accatgtcct	ttcaacttga	aaattctatg	attctgtata	atcaaacaac	8280
ctaactcact	ggatTTTca	ctttcaacc	agctattccg	tgtcttacag	tgctctggtc	8340
ttcacagata	ggtaggcacc	tggcatgtca	gttgcctgaa	tttgcggat	ttcacctgtta	8400
tgttttggta	cgataaaaaat	aaaaatgtaa	tttatataatc	tgaatcaggt	ctgtatgtta	8460
tgatcaattg	ctcagcaatt	tcgggcagtt	ggtttgatgg	ttatgttagta	atgttagcctg	8520
agagcagaaa	tacagagcct	ctggggctaga	gaaagtataa	atggcatcct	aggctatgtta	8580
gggttacagc	tcttcagaag	gaactttcat	tttcattgtg	acacatcgtc	tacatgttgt	8640
agaagaacat	agtttcagaa	tttccatcagt	tagaaacata	gtttctcaaa	tattcacttt	8700
cagcattggg	tagaaaaagt	acccaggtga	tggactgtatt	gaggatccac	tagtaacggc	8760
cggccatgtg	cttggaaattct	gcagatatcc	atcacactgg	cggccgctcg	agtctagagg	8820
gcccgtttaa	acccgctgat	cagcctcgac	tgtgccttct	agttgccagc	catctgttgt	8880
ttggccctcc	cccggtgcctt	ccttgcacct	ggaagggtgcc	actcccactg	tcctttccct	8940
ataaaaatgag	gaaattgcat	cgcattgtct	g			8971

```
<210> 43
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 43
cggatccccq qcgcatacgt caactgtct
```

```
<210> 44
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 44
taaaagcttgc caccatgaag ctactgtctt
```

```
<210> 45
<211> 10514
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant
      DNA vector sequence

<400> 45
ctaaattgtt aagcggttaata ttttggtaaaa attcgcgtt aatttttgtt aaatcagctc 60
attttttaac caataggccg aaatcgccaa aatcccttat aaatcaaaag aatagaccga 120
gatagggtt agtgttgttc cagtttggaa caagagtcca ctattaaaaga acgtggactc 180
caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactaacgtg aaccatcacc 240
ctaattcaagt ttttgggtt cgaggttgcgg taaagcacta aatcgaacc ctaaaggag 300
cccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaaagg aagggaaagaa 360
aqcqaaaqqa qcqqqcqcta qqqcqctqgc aqgtgtaqcq gtcacqctgc qcgttaaccac 420
```

cacacccgccc	gcgcttaatg	cgccgctaca	gggcgcgtcc	cattcgccat	tcaggctgcg	480
caactgttgg	gaagggcgat	cggtgcgggc	ctcttcgcta	ttacgcgcagc	tggcgaaagg	540
gggatgtgct	gcaaggcgat	taagttgggt	aacgccaggg	tttcccagt	cacgacgttg	600
aaaaacgacg	gcoagtgagc	gcccgtata	cgactcacta	tagggcgaat	tggagctcca	660
ccgcgggtcg	gccccggcat	gcaggccaga	catgataaga	tacattgtat	agtttggaca	720
aaccacaact	agaatgcagt	aaaaaaaaatg	cttattttgt	gaaatttgg	atgttattgc	780
tttatttgta	accattataa	gctcaataa	acaagttgtc	gagttacttg	tacagctcg	840
ccatgccgag	agtgatccccg	gcccgtca	cgaactccag	caggccatg	tgatcgcgct	900
tctcggtgg	gtctttgtctc	agggcggact	ggtagctcg	gtagtgg	tgccggcagca	960
gcacggggcc	gtcgcgcgtat	gggtgttct	gctggtagtg	gtcggcggc	tgcacgctgc	1020
cgtcctcgat	gttggggcgg	atcttgaagt	tcaccttgc	gccgttctc	tgcttgcgg	1080
ccatgatata	gacgttgtgg	ctgtttagt	tgtactccag	cttggcccc	aggtatgtgc	1140
cgtcctcctt	gaagtcgtat	cccttcagct	cgatgcgggt	caccagggtg	tcgcccctcg	1200
acttcacctc	ggcgcgggtc	ttgttagtgc	cgtcgtcctt	gaagaagatg	gtgcgcctc	1260
ggacgtagcc	ttcgggcatg	gcccgttgc	agaagtcgtg	ctgcttcatg	tgttcgggt	1320
agcggcgaa	gcactgcagg	ccgtagccga	aggtgtc	gagggtggc	cagggcacgg	1380
gcagcttgc	gggtgtcag	atgaacttca	gggtcagctt	gccgttagtg	gcatgcgcct	1440
cgccctcgcc	ggacacgcgt	aacttgcgc	cggttacgtc	gccgtccagc	tcgaccagga	1500
tgggcaccac	cccggtgaac	agtcctcgc	cctgctcac	catggtggc	ttagcttgc	1560
ttgacagtgg	ctgggggttg	cgccgcggg	tttata	agccacagcg	gcccactcg	1620
ccataaaaagg	caactttagg	aacggcgggg	ggtgatttgg	ttcgagtcgt	ttatttcacccg	1680
gccttgcgc	acagtgcagc	attttttac	ccctctccc	ctcctttgc	gggggaaaaaa	1740
aaaaaaaaaa	aaaaaaaaaaag	gagaagagaa	aaaaagcgag	cgagagagaa	agcgagattg	1800
aggaagagga	tgaagagttt	ggcgatgggt	gctgg	ttccag	tggacaagaa	1860
tagccccgc	ccttgcggac	agtatccat	tcagtgcact	agatcagtca	agcggccgcg	1920
tctaacctag	gagccaccca	cagacactat	actataacca	cactgaattt	tcttccttat	1980
ctcaaataat	cctattttgt	ttcatgcctc	atgtgcctt	aacaagcaga	tttctcgcc	2040
agacatgggt	gcttgcgcct	gtaatcccag	cacttggcc	aagtccggca	gtcacttgc	2100
ggtcaggagt	ttgaatttgc	cctgtccat	atgtgaaac	cccatctct	ctaaaaataac	2160
aaaaatttagc	tggcatgtt	ggtgtgc	tgttagtccc	gtacttcagg	agactgaggc	2220
aggagaatttgc	cttgaactcg	ggaagcagag	gttgcagtga	gccaagatca	tgccattgc	2280
ctccagcctg	ggcaacagag	cgagactccg	tctcaa	aacaaaacaa	aaacaaaaaa	2340
acaagcagat	ctotcaactt	ggaatgtgc	tgcttctca	tgccaa	cttttttttt	2400
tttcttttag	actcaactca	aatatgatcc	tccctgaagt	ctctccatct	ttccttagaca	2460
caattagcaa	ccagttcctc	tgtgttcaa	cagcacttcg	tgtgtatgt	tatgtatgt	2520
tatttatcat	tattattatc	accttgc	actctccctg	aaggtaatag	ccaggcactt	2580
agtatagagt	aggtatcaac	aattgtt	tgaaggata	aatgggtgaa	tgacaacttt	2640
gcaaaggttc	aatttttaat	ctattgattt	gaaagaatta	ttatacatta	ttttaaatat	2700
ccttcttagt	tctaataatct	tatgtctata	gtatgtcgac	agttacttta	aatcagatata	2760
taaatgattt	tcattttaca	aagaacagac	ctgccttct	caatgtcaca	agactattgt	2820
agagaatagc	agtgtaaataa	gtacatgtga	aaatacttgg	caaagtaata	gtgaaaataag	2880
agctgatctc	acaggaatgt	tttggaaattt	aaaacaccgt	atgtaaaata	ggtatctaa	2940
tctgtggc	catattttaa	ctctcaataa	atattaactt	tgattatttta	aaaaatacct	3000
ggtagctttt	ttgtgaattc	tacaagaacc	tgtacatgt	tggttgccat	ttccgtcaaa	3060
atgagaaaat	tttcttctgc	actgaatttct	tcttttact	aaatttttaag	aaaaaaacca	3120
atgattaaaa	agttagataag	aaaatgataa	tataaaacca	tttagactaaa	aaaaaaactag	3180
gtttagatga	aaaaatttga	aaacccatct	tactatttct	cagatcttgc	gagggtcaatt	3240
accattatca	agatagggtgg	gagactctat	actcttgc	tgtttatgt	tgctgagat	3300
taaaaagtct	aacttgcattt	ctcagctcta	atcatttca	agtaacatata	tccatataaa	3360
attatactc	caaggcaata	aaaactccag	aaaatattaa	ctttaaaccc	acatattttc	3420
agatattata	catacaattt	tatttttat	ttcctgaggc	atcctctgtt	tgttatatga	3480
atccataata	aaatgttagaa	gagtctgttgc	atctgggttgc	agttcgtttt	tccctacac	3540
tggcaagaaa	ataaaagggtt	ttagggaaaga	tgtgtataaa	gccttcttc	tatgtgtca	3600
ttcatcttgc	gggacctaatt	agcagaaatc	ataaaaaaaa	aaaggcataa	atacagcagg	3660
aactatgagt	tctaagagag	tgtataattt	cttctcttgc	gtcctcaggc	agccatatac	3720
acatttcaact	gaattttctc	atctataaaa	tatgtataac	aaaatattcc	acttgcgtat	3780
ttcttaggaa	taaatggtag	tctggatata	agactatttgc	agcccattca	tattgtgtat	3840
cactgaaatg	tcaagtttca	attataacta	atgaaagggtt	ttcataatgt	atgactacta	3900

tgaatactat	attcatatac	tattctacat	gaattccacc	tggtaagtc	atctgagagt	3960
attcatgtgg	aggagggatc	aacaaactgt	ggcccatgga	ccaaatccat	tggattttat	4020
aaagtttat	tggaacacag	tcacactcat	ttgttatgt	attgtctctg	ctactttgc	4080
actaaaatgg	cagagtttag	tagctttgt	agataccatg	tggccaaagc	cctttaaaga	4140
aaaaaagttt	gtcaacccct	aaaatagaga	gtacaacaatc	tggacaattc	agaccccaa	4200
catctcttc	tccctctaca	gcatctttg	aagagtgtga	ataattatgt	caactcttg	4260
ttttcttata	ccaaagtgtt	actgatgaag	cactatggca	aaataatgtc	ttcatataaa	4320
atacttttc	ttaataataat	acaaaataac	gtatttgtaa	aactaaatgt	ggataaagt	4380
ttctttgtca	ataatgtaat	aaatattgag	aagggtgtgg	tttggacaat	catagttac	4440
tggttttca	gctaattggca	aaataaaaatt	ccatcaaatg	tcctgtgtcat	tatcacaatg	4500
tgtgtgtcta	aacaccttgg	atgctttcc	cctattttg	cctaagttgc	tcaatggatt	4560
tcttgagaaa	tttgtaaccc	cagtggtag	gacttttctc	tattcatcca	catacggtta	4620
ttcacccctgt	tttactcatt	tataaggatt	aatccacagc	ctctatttgc	accatgttct	4680
tcccagttt	agactgtctt	ctcaaaaaga	ggaaaaaaat	ggtcatcaat	actcagttgg	4740
aaaaaatcca	tttatacac	caaattcatt	ctgctggcaa	acactaggt	gtaggactgg	4800
aatggagtca	gagagaagga	aaacttcac	tttttagttt	atgcattgt	actatttga	4860
tgtttacatc	aaagcatgt	ttaggaatcc	attnaatga	acaattggac	tagaaattca	4920
gctgtgcctg	tggcttaat	ttgaaagcct	cctggccagc	tcactatgt	cctagaaaact	4980
cagttcctgc	cagtcttggc	agacaccatt	gcatattatt	accctgcatt	actttgttgc	5040
cgaggtcact	tgtcgcaagt	cacgaccc	actgtcttct	ctagacggcg	atacagtcaa	5100
ctgtcttgc	ccttgcattac	tactcttcc	cgatgatgt	gtcgactta	ttctatgtctg	5160
tctcaatgtt	agaggcatat	cagtctccac	tgaagccat	ctatctgt	cggcatctt	5220
attcacatta	tcttgacaa	ataatcctgt	taacaatgt	tttataatcct	gtaaagaatc	5280
catttcaaa	atcatgtcaa	ggtcttc	aggaaaaatc	agtagaaata	gtgttccag	5340
tctttcttagc	tttgattcca	cttctgtcag	atgtgcctca	gtcagcggag	actttttgg	5400
tttgggagag	tagcgacact	cccagtgtt	cttcagacac	ttggcgact	ttgggttttc	5460
tttggagcac	tttagctttt	taagtccgca	aatatcgcat	gcttgcgtga	tagaagacag	5520
tagcttcatg	gtggcgatc	ccgataagcc	agtaaggcgt	gggttctct	gttagccaga	5580
gagctctgt	tatatagacc	teccacctga	cacgcctacc	gccccatgtc	gtcaatgggg	5640
cggagtttgc	acgacattt	ggaaagtc	gttgcatttt	gtgcwwwaa	aaactccat	5700
tgacgtcaat	gggggtggaga	cttggaaatc	cccgtgagtc	aaaccgctat	ccacgcccatt	5760
tgatgtactg	ccaaaaccgc	atcaccatgg	taatagcgat	gactaatacg	tagatgtact	5820
gccaagtagg	aaagtcccat	aaggtcatgt	actgggcata	atgccaggcg	ggccattttac	5880
cgtcattgac	gtcaataggg	ggcgtacttg	gcataatgata	cacttgatgt	actgccaagt	5940
gggcagttt	ccgtaaatac	tccaccatt	gacgtcaat	gaaagtccct	attggcgta	6000
ctatggAAC	atacgtcatt	attgacgtca	atgggcgggg	gtcgttggc	gtcagccag	6060
gccccccatt	taccgtta	tgatgtacgc	ggaactccat	atatggct	tgaacta	6120
accccgtaat	tgattactat	taataactag	tcaataatca	atgtcaacgc	gtatatctgg	6180
cccgtaatc	gcgaaggcgc	gcaaaacg	taaccctacc	cgggctgcag	gaattctacc	6240
gggttagggg	ggcgctttc	ccaaggcagt	ctggagcat	cgctttagca	gccccgctgg	6300
cacttggcgc	atcacaagt	gcctctggc	tcgcacacat	tccacatcca	ccggtagcgc	6360
caaccggctc	ccttcttgg	tggcccttc	gcccac	ctactcc	cctagtctgg	6420
aagttcccc	ccgcccccgca	gctcgctcg	tgaggacgt	gacaaatgg	atagcacgt	6480
ctcaactagtc	tcgtcgat	ggacaagcac	cgctgagcaa	tggagcggg	taggccttt	6540
gggcagcggc	caatagcgc	ttggctc	cgcttctgg	gctcagaggc	tgggaagggg	6600
tgggtccggg	ggcggtctca	ggggcggct	cagggcgggg	gccccgcgc	aggtctcccg	6660
gaccggccat	tctgcacgt	tcaaaagcgc	acgtctgc	cgctgttctc	ctcttctctca	6720
tctccggcc	ttcgacatgc	atgaaaaaagc	ctgaactc	cgccgacgt	gtcgagaagt	6780
ttctgtatca	aaagttcgac	agcgtctccg	actgtatgc	gtctcggag	ggcgaagaat	6840
ctcggttt	cagttcgat	gttaggggc	gttgcattgt	cctgcgggt	aatagctcg	6900
ccgatggtt	ctacaaagat	cgttat	atccggactt	tgcatcggcc	gogctcccg	6960
ttccggaaat	gtttgcatt	gggaaatca	gcgagagcct	gacctattgc	atctcccgcc	7020
gtgcacaggg	tgtcacgtt	caagacctgc	ctgaaaccga	actgcccgt	gttctgcagc	7080
cggtcgccg	ggccatggat	gcatcgctg	cgcccgatct	tagccagacg	agcgggttc	7140
gcccattcgg	accgcaagga	atcggtcaat	acactacat	gcgtgat	ttcatacgca	7200
ttgctgtatcc	ccatgtgtat	cactggcaaa	ctgtatgg	cgacaccgtc	atgtcg	7260
tcgcgcaggc	tctcgatgag	ctgatgttt	ggcccgagga	ctgccccgaa	gtccggcacc	7320
tcgtgcacgc	ggatttcggc	tccaacaat	tccgtacgg	aatggccgc	ataaacagcgg	7380

tcattgactg gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct	7440
tctggaggcc gtgggtggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc	7500
cggagcttgc aggatcgccg cggctccggg cgtatatgct ccgcatttgtt ctgaccaac	7560
tctatcagag cttgggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg	7620
acgcaatcgt ccgatccgga gccgggactg tcgggcgtac acaaattcgcc cgacagaagcg	7680
cggccgtctg gaccgatggc tggtagaaag tactcgccga tagtgaaaac cgacgccccca	7740
gcactcgtcc gagggcaaag gaatagagta tgccagaaatt gatgatctat taaaataaa	7800
agatgtccac taaaatggaa gttttctg tcatactttg ttaagaagg tgagaacaga	7860
gtacctacat ttgtaatggg aggattggg ctacgggggtt ggggggggg tgggattaga	7920
taaatgcctg ctcttactt aaggctttt actattgtt tatgataatg tttcatagtt	7980
ggatatacata atttaaacaa gcaaaaccaa attaaggccc agtcattcc tccactcat	8040
gatctataga tctatagatc tctcggtggg tcttgcattt tctcttgatt cccactttgt	8100
ggttctaagt actgtggttt ccaaattgtt cagtttccata gcctgaagaa cgagatcagc	8160
agcctctgtt ccacatacac ttcatttca gtattgttt gccaaggctt aattccatca	8220
gattttatgt gacatctta ttgttataa gatcatcaat ttctgcatac taagcttac	8280
gataccgtcg acctcgaggg gggcccggt acccagctt tgccctt tttttttt agtgagggtt	8340
aattgcgcgc ttggcgtaat catggcata gctgtttcct gtgtaaaatt gttatccgct	8400
cacaattcca cacaacatac gagccggaag cataaaagtgt aaagccctggg gtgcctaatg	8460
agttagctaa ctacacattaa ttgcgttgcg ctcaactgccc gctttccagt cggaaaccc	8520
gtcgtgccag ctgcattaat gaatcgccca acgcgcgggg agaggccgtt tgcgtattgg	8580
gcgcctttcc gcttcctcgc tcactgactc gctgcgtcg gtcgttgcg tgcggcgagc	8640
ggtatcagct cactcaaagg cgtaatacg gtatccaca gaatcagggg ataacgcagg	8700
aaagaacatg ttagaaaaaa gcccaggaac cgtaaaaagg cccgcgttgc	8760
ggcgttttc cataggctcc gccccctga cgagcatcac aaaaatcgac gtcagaatca	8820
gagggtggcga aacccgacag gactataaag ataccaggcg tttcccccgtt gaagctccct	8880
cgtgcgtctt cctgttccga ccctgcgtt tacggatac ctgtccgcct ttctcccttc	8940
gggaagcgtg ggcgttttc atagctcacg ctgttaggtt ctcaatgggg ttaggtcg	9000
tcgcctcaag ctgggtgttgc tgacgaaacc cccgcgttgc cccgaccgtt ggccttattc	9060
cggtaactat cgttttgat ccaaccgggtt aagacacgac ttatcgccac tggcagcagc	9120
cactggtaac aggattagca gagcgaggta ttagggcggt gctacagat ttttggatgt	9180
gtggcctaac tacggctaca cttagggac agatattttt atctgcgttc tgctgaagcc	9240
aggtaaccttcc gggaaaaagag ttgttagctc ttatccggc aaacaaacca cccgtggtag	9300
cggtgggtttttttt ttttttttca agcagcagat tacgcgcaga aaaaaggat ctcagaaga	9360
tcctttgtatc ttttctacgg ggtctgacgc tcagtggaaac gaaaacttcac gtaaggat	9420
tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag	9480
ttttaaatca atctaaatgtatgata aacttggtct gacagttaacc aatgcttaat	9540
cagtggggca cctatctcag cgatctgtct atttcgttca tccatagttt cctgactccc	9600
cgtcgtgttag ataactacga tacggggaggctt accatctt ggcggccgtt ctgcaatgt	9660
accgcgagac ccacgctcac cggctccaga ttatcagca ataaaccagc cagccggaaag	9720
ggccgagcgc agaagtggtc ctgcaactttt atccgcctcc atccagtctt ttaatttttgc	9780
ccggaaagct agagtaagta gttcgccagt taatagtttgc caacacgttgc ttggcattgc	9840
tacaggcatc gtgggtgtcac gtcgtcggtt tggatggct tcattcagct cccgttccca	9900
acgatcaagg cgagttacat gatccccat ttgtgcaaa aaagcggtt gtccttcgc	9960
tcctccgatc gttgtcagaa gtaagtggc cgacgtgttca tcactcatgg ttatggcagc	10020
actgcataat tctttacttgc tcatgcacatc cgttaagatgc ttttctgttgc ctggtgagta	10080
ctcaaccaag tcattctgag aatagtttat gggcgaccg agttgtctt gcccggcggtc	10140
aatacgggat aataccgcgc cacatagcag aactttaaaat gtgtcatca ttggaaaacg	10200
ttcttcgggg cggaaaactctt caaggatctt accgctgttg agatccagtt cgatgtaaacc	10260
cactcggtca cccaaactgtat cttcagcatc ttttactttc accagcggtt ctgggtgagc	10320
aaaaacagga aggcaaaaatg ccgcaaaaaa ggaataagg ggcacacggaa aatgttgaat	10380
actcataacttcc ttcttttca aatattatttgc aagcattttt cagggttatt gtctcatgag	10440
cggatacata ttttagatgtt ttttagaaaaaa taaacaaata ggggttccgc gcacatttcc	10500
ccggaaaatgttccgc ccac	10514

<210> 46
<211> 9626
<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic recombinant DNA vector sequence

<400> 46

gacggatccgg gagatctccc gatccccat ggtcgactct cagtacaatc tgctctgatg
ccgcatagtt aaggcagtat ctgctccctg ctgtgtgtt ggaggtcgct gagtagtgcg
cgagcaaaat ttaagctaca acaaggcaag gcttgcaccga caattgcatt aagaatctgc
ttagggttag gcgtttgctg atgcttcgcg atgtacgggc cagatataacg cgttgacatt
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata
tggagttccg cgttacataa cttacgtaa atggcccgcc tggctgaccg cccaacgacc
cccgcccatt gacgtcaata atgacgtatg ttcccatatgt aacgccaata gggactttcc
attgacgtca atgggtggac tatttacggt aaactgccc cttggcagta catcaagtgt
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt
atgcccagta catgacccctt tgggactttc ctacttggca gtacatctac gtattagtca
tcgctattac catggtgatg cgggtttggc agtacatcaa tgggcgtgga tagcggtttgc
actcacgggg atttccaagt cttccacccca ttgacgtcaa tgggagtttgc ttttggcacc
aaaatcaacg ggacttccaa aatgtcgta acaaactccgc cccattgacg caaatgggc
gttaggcgtgt acggtggggag ctgtctactg gtcataataa gcagagctct ctggctaact
gatggacttccaaatc aagttggatgtt agataatccag cacaggccca actggagcc
tgcggcccg aaggaagata gccaaggctt actgaggctg
gttggagggag ccactgctgg gtcgtgttgc gcccgttacc
tgggctccag atttgggagc tctctgtct tagagtccct
gacactgggg ctcctccct tgcgtgttgc
tgcttataac atgtccttcc
ggcctggctg tgccctctggg
attggtgagg caaagtccagg
gtgcacctgt gatgcttcca
tatgactccc tgagccactc
aaggaatgga acctaaggag
tccatcatct tctcacatgt
atgggctaga gcgctggagc
agtttcccaag taactttcca
cacaagggtc tggatgggt
tgctgccccct cccgagcacc
atccctggacg catgaaaacta
cttttatctc ttttgttaaa
ggaatctgtat ttgtcaaggg
cctgtctgcc atctgtgtgc
gatctgggtt tggccattca
ttgccagtgt gggcatcatc
gcctgggggtt ccagggcgat
tgggggggtt actgggagcc
tggataaattc ttagaacgtat
ccttacatct ggccttatgc
cctccttctc ttagaggtgt
ctggcatctt ttggccctcag
gttggggctt gccagactgt
acagcccagg gtcataatata
ctccctgtct tccttccccg
aacggggctc tgccctgaga
ctgagtccgg ggcacccacc
ccctqqqqcc ttgcagggcc


```

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 47
gagaaccac tgcttactgg ct 22

<210> 48
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 48
gcattgtgt gcagtctgt a 21

<210> 49
<211> 7354
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant
      DNA vector sequence

<400> 49
cgcgattgct gatccccatg ttttatcaactg gcaaaactgtg atggacgaca ccgtcagtgc 60
gtccgtcgcg caggctctcg atgagctgtat gctttgggcc gaggactgccc cccaaagtccg 120
gcacccctgtg cacgcggatt tcggctccaa caatgtccctg acggacaatg gcccataaac 180
agcgggtcatt gactggagcg aggcgtatggtt cggggattcc caatacgagg tcgccaacat 240
cttcttcgttgg aggccgtggt tggcttgtat ggagcagcag acgcgtact tcgagcggag 300
gcataccggag cttgcaggat cgccgcggct ccgggcgtat atgcctccga ttggcttta 360
ccaaactctat cagagcttgg ttgacggcaa ttgcgtatgtat gcagcttggg cgcagggtcg 420
atgcgacgca atcgccat ccggagccgg gactgtccgg cgtagcacaaa tcgcccgcag 480
aaggcgccgc gtctggaccg atggctgtgt agaagactc gcccataatgt gaaaccgacg 540
ccccacgact cgtcccgaggg caaaggaata gaggatgcag aaattgtatgtat tctattaaac 600
aataaaagatg tccactaaaa ttggaaatttt tccgtcata ctttgttaag aagggtgaga 660
acagagttacc tacatTTGA atggaaaggat ttggactacg ggggtggggg tgggggtggga 720
tttagataaaat gcctgtctt tactgaaggc tctttactat tgctttatgtat taatgtttca 780
tagttggata tcataattta aacaagcaaa accaaattaa gggccagctc attcctccca 840
ctcatgtatct atagatctat agatctctcg ttggatcatt gttttctct tgattccac 900
tttgtgggttc taagtactgt gggttccaaa ttgtgtcaggat tcatagccctg aagaacgaga 960
tcagcagccct ctgttccaca tacacttcat tctcaggat gttttgcctt gttctaaattc 1020
catcagattt tagtgacat cttttatgtt taatagatca tcaatttctg catabacta 1080
ttatcgatac cgtcgaccctc ggggggggc ccggatccca gcttttgcctt cctttatgtga 1140
gggttaatttgcg cgcgttggc gtaatcatgg ttcatatgtt gtttttttttgcgtt 1200
ccgcctcacaat ttccacacaa catacgagcc ggaaggataaa agtgtaaagc ctgggggtgc 1260
taatgtatgtat gctaactcattaaatgtcg ttgcgtcact tggccgtttt ccagtcggg 1320
aacctgtcgat gccagctgca ttaatgtatgtat gggcaacgcg cggggagagg cgggttgcgt 1380
attggggcgctt ctggcgttcc ctcgtcact gactgtcgat gctcggtcgat tcggctgcgg 1440
cgagcggat cagctcactt aaaggccgtt atacggatccat ccacagaatc aggggataac 1500
gcaggaaaga acatgtgagc aaaaggccag caaaaggccca ggaaccgtaa aaaggcccg 1560
ttgtggcgt tttccatag gctccggccc cctgacgagc atcacaataaa tcgacgctca 1620
agtcaagatgg ggcgaaaccc gacaggacta taaagatacc aggcgttcc ccctggaaagc 1680

```

tccctcggtc	gctctcctgt	tccgaccctg	ccgcttacccg	gataacctgtc	cgcccttc	1740
ccttcggaa	gcgtggcgct	ttctcatagc	tcacgctgt	ggtatctcg	ttcgggtgt	1800
gtcggtcgct	ccaagctggg	ctgtgtcac	gaacccccc	ttcagccga	ccgctgcg	1860
ttatccgt	actatcgct	tgagtccaa	ccggtaagac	acgacttac	gccactgg	1920
gcagccactg	gtaacaggat	tagcagagcg	agttatgt	gccccgtac	agagttctt	1980
aagtggggc	ctaactacgg	ctacactaga	agacagat	tttgtatct	cgctctg	2040
aagccagtt	ccttcggaaa	aagagtgg	agctctt	ccggcaaaca	aaccaccg	2100
ggtagccgt	gtttttgt	tttcaagc	cagattac	gcagaaaaaa	aggatct	2160
gaagatctt	tgtatctt	tacgggg	gacgct	ggaacaaaaa	ctcacgtt	2220
gggatgg	tcatgagatt	atcaaaaa	atcttcc	agatcc	aaataaaaaa	2280
tgaagttt	aatcaatcta	aagtatata	gagtaaa	ggctgac	ttaccaat	2340
ttaatcg	aggcacct	ctcagc	tgtctat	gttcatcc	agttgc	2400
ctcccg	tgtagataac	tacgata	gaggg	catctgg	cagtgc	2460
atgataccgc	gagacccac	ctcacc	ccagat	cagcaataa	ccagcc	2520
ggaaggccg	agcgcaga	tggtc	actttat	cctccat	gtctat	2580
tgttgcggg	aagctag	aagtat	ccagtt	gttgcg	cgttgt	2640
attgtacag	gcacgt	gtcacg	tcgtt	tggctt	cagctcc	2700
tcccaacgat	caaggcg	tacatgat	ccc	gcaaaaaa	ggttag	2760
tccgttctc	cgatcg	cagaag	ttggcc	tgttac	catgg	2820
gcagcactg	ataattct	tactgt	ccatcc	gatgc	tgtact	2880
gagactcaa	ccaagt	ctgagaat	tgtat	gaccg	ctcttgc	2940
gcgtcaatac	gggataatac	cgcg	acat	taaaagt	catcatt	3000
aaacgttctt	cggggc	actct	atcttac	tgtt	gagat	3060
taacccactc	gtgcac	ctgat	ccat	tttca	cgttct	3120
tgagcaaaaa	caggaagg	aaatgc	aaaaagg	taagg	gcac	3180
tgaataactc	tact	tttcaat	tatt	gaa	tttac	3240
atgagccgat	acat	tattt	atgtat	aaaataa	aatagg	3300
tttcccgaa	aagtgc	taaatt	gtt	taat	ttcgcgtt	3360
attttgtt	aatcag	ttttttaa	aatagg	ccaa	atccctt	3420
aatcaaaaga	atagacc	atagg	gtt	tttcc	agtgatcc	3480
tattaaagaa	cgtggact	aacgt	ggc	aaaaaa	cgttcat	3540
caactacgt	accat	taat	tttgg	gagg	tgccgt	3600
atcggaaccc	taaagg	cccc	gat	gggg	ggccagg	3660
cgagaaagga	agg	gat	tttgc	gggg	gcaacgt	3720
tcacgctcg	cgtaacc	acacc	ccg	cttgc	ggcg	3780
attcgccatt	cagg	actgt	ggg	tttgc	ggcg	3840
tacgcca	ggc	ggat	gtc	tttgc	tttgc	3900
tttcccagtc	acgac	aaa	ccag	tttgc	tttgc	3960
agggcgaatt	ggag	ctcgtt	ggg	tttgc	tttgc	4020
acattgat	gttgg	accaca	act	tttgc	tttgc	4080
aaatttgt	tgtt	tttgc	tttgc	tttgc	tttgc	4140
agttactt	acag	tttgc	tttgc	tttgc	tttgc	4200
aggaccat	gtcgc	tttgc	tttgc	tttgc	tttgc	4260
tagtgg	cgggc	cacgg	tttgc	tttgc	tttgc	4320
tccggag	gcac	gtcct	tttgc	tttgc	tttgc	4380
ccgttctt	gttgc	catgat	tttgc	tttgc	tttgc	4440
ttgtgc	ggat	tttgc	tttgc	tttgc	tttgc	4500
accagggt	cgc	tttgc	tttgc	tttgc	tttgc	4560
aagaagat	tgc	tttgc	tttgc	tttgc	tttgc	4620
tgctt	gttgc	tttgc	tttgc	tttgc	tttgc	4680
agggtgg	agg	tttgc	tttgc	tttgc	tttgc	4740
ccgttag	catgc	tttgc	tttgc	tttgc	tttgc	4800
ccgtcc	cgacc	tttgc	tttgc	tttgc	tttgc	4860
atgg	tttgc	tttgc	tttgc	tttgc	tttgc	4920
gccac	tttgc	tttgc	tttgc	tttgc	tttgc	4980
tcgag	tttgc	tttgc	tttgc	tttgc	tttgc	5040
tcctt	tttgc	tttgc	tttgc	tttgc	tttgc	5100
gagag	tttgc	tttgc	tttgc	tttgc	tttgc	5160

aggcccagat ggacaagaat agcccccgcc cttgcggaca gtagccatt cagtactca	5220
gatcgtcaa gcggccgctc tagacggcga tacagtcaac tgcgttgc ctttgttact	5280
actctttcc gatgtatgat tcgcacttat tctatgtgt ctcaatgtta gaggcatatc	5340
agtctccact gaagccaatc tatctgtgac ggcatctta ttcacattat cttgtacaaa	5400
taatccgtt aacaatgtt ttatatccgt taaaagaaatcc attttcaaaa tcgttcaag	5460
gtcttcgtca gaaaaaatca gtagaaatag ctgtccagt ctttctagcc ttgattccac	5520
ttctgtcaga tggcccttag tcagcggaga cttttgggtt ttgggagagt agcgacactc	5580
ccagttttc ttcaagacact tggcgcaatt cggttttct ttggagact tgagctttt	5640
aagtgcgaa atatcgatg ctgttcgtat agaagacagt agtttcatgg tggcgatcc	5700
cgataagcca gtaaggcgtg gttctctag tttagccagag agtctgtt atatagacct	5760
cccaccgtac acgcctaccg cccatttgcg tcaatggggc ggagttgtta cgacattttg	5820
gaaagtcccg ttgattttgg tgccaaaaca aactccatt gacgtcaatg ggtggagac	5880
ttggaaatcc ccgtgagtca aaccgtatc caegcccatt gatgtactgc caaaaccgca	5940
tcaccatggt aatagcgatg actaatacgt agatgtactg ccaagtagga aagtcccata	6000
aggtcatgtt ctggcataa tgccaggcgg gccatttacc gtcatttacg tcaatagggg	6060
gcgtacttgg catatgatac acttgatgtt ctgccaagtg ggcagttac cgtaaataact	6120
ccaccoattt acgtcaatgg aaagtcccta ttggcggttac tatggaaaca tacgtcatta	6180
ttgacgtcaa tggcggtttt tcggtggcgt gtcagccagg cgggcccattt accgttaagt	6240
atgttaacgcg gaaactccata tatggctat gaactaatga ccccgtaatt gattactatt	6300
aataactgtt caataatcaa tgtcaacgcg tatatctggc cgcgtacatcg cgaaggcagcg	6360
caaaaacgcct aaccctaccc gggctgcagg aattctaccg ggttagggag ggcgttttcc	6420
caaggcagtc tgagcatgc gcttttagcag ccccgctggc acttggcgca tcacaagtgg	6480
cctctggctt cgcacacatt ccacatccac cgttagcgcc aaccggctcc cttttttgg	6540
ggcccccctcg cggccacccctt tactccctcc cttagtcagga agttcccccc cggcccccag	6600
ctcgctgtgc gcaaggcgtg acaaattggaa gtgcacgtc tcaactgtt cgtcagatg	6660
gacaaggcacc gctgagcaat ggaagcggtt aggcctttgg ggcagccggc aatagcagct	6720
tggctcttc gctttctggg ctcagaggct ggaagggtt gggccgggg gggccgtcag	6780
ggccgggctc aggggcgggg cggcgccgaa ggtccctccgg accggcatt ctgcacgtt	6840
caaaaacgcga cgtctgccgc gctgttctcc tttccctcat ctccggccct tcgacctgca	6900
tgaaaaaagcc tgaactcacc ggcacgtctg tcgagaaggat tctgatcgaa aagttcgaca	6960
gcgtctccga cctgatgcag ctctcgagg gcaagaatc tcgtgttcc agttcgatg	7020
taggagggcg tggatgttc ctgcgggtt atagctgcgc cgatggttt tacaagatc	7080
gttatgttta tcggcactt gcatcgccg cgctcccgat tccggaaatg ctgcacattt	7140
gggaattcag cgagagcctg acctattgc tctcccgcc tgcacagggt gtcacgttgc	7200
aagacactgc tggaaacggaa ctgcccgtt ttctgcagcc ggtcgccgg gccatggatg	7260
cgatcgctc ggccgatctt agccagacga gcccgggttccg cccattcgga ccgcaaggaa	7320
tcggtaata cactacatgg cgtgatttca tatg	7354

<210> 50

<211> 10514

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic recombinant
DNA vector sequence

<400> 50

ctaaatttgcgta agcgtaata tttttttaaa attcggttta aattttttttt aaatcagctc	60
attttttaac caataggccg aaatccgaa aatcccttat aaatcaaaag aatagaccga	120
gatagggttgc agtgggttc cagttggaa caagagtcca ctatcaaaga acgtggactc	180
caacgtcaaa gggcgaaaaa cgcgttatca gggcgatggc ccactacgtg aaccatcacc	240
ctaatcaatgtt tttttgggtt cgagggtccg taaagcacta aatcggaaacc ctaaaggggag	300
cccccgattt agagtttgcg gggggaaagcc ggcgaacgtg gcgagaaagg aagggaaagaa	360
agcggaaagga gccccggctt gggcgctggc aagtgttagcg gtcacgttgc gcttaaccac	420
cacacccggcc gcgcttaatg cgcgttacg gggcgctcc cattcccat tcaggctgcg	480
caactgttgg gaaaggcgat cgggtcgccg ctcttcgttta acggccagc tggcgaaagg	540

gggatgtgct	gcaaggcgat	taagttgggt	aacgccaggg	tttcccagt	cacgacgtt	600
taaaacgacg	gccagtgagc	gcgcgtata	cgactca	tagggcgaat	tggagctcca	660
ccgcggcg	gccgggcat	gcaggccaga	catgataaga	tacattgatg	agtttgaca	720
aaccacaact	agaatgcagt	aaaaaaaatg	cttatttg	gaaatttgc	atgttattgc	780
tttatttgc	accattataa	gctgcaataa	acaagttgtc	gagttacttg	tacagctcg	840
ccatgccgag	agtatccccg	gccccggtca	cgaactccag	caggaccatg	tgatcgcgct	900
tctcggtgg	gtcttgc	agggcgact	gtagtctcg	gtagtggtt	tcgggcagca	960
gcacggggcc	gtcggcgatg	gggggttct	gctggtagt	gtcggcgagc	tgcacgctgc	1020
cgtcctcgat	gttgtggcg	atcttgc	tcacccgt	gccgttctc	tgcttgcgg	1080
ccatgatata	gacgttgtgg	ctgttgtgt	tgtactccag	cttgcgtcccc	aggatgttgc	1140
cgtcctcctt	gaagtgcgt	cccttcagct	cgtgcgggtt	caccagggtg	tcgcctcg	1200
acttcaccc	ggcgccggtc	ttgttgtgc	cgtcgtcctt	gaagaagatg	gtgcgtcct	1260
ggacgtagcc	ttcgggc	catg	gcggacttga	agaagtcgt	ctgcttc	1320
agcggggc	gactgcagg	ccgtagccga	agttgtcac	gaggggtggc	cagggcacgg	1380
gcagctg	gggtgtcag	atgaacttca	gggtcagctt	gccgttaggt	gcatcgccc	1440
cgccctcgcc	ggacacgctg	aacttgtggc	cgttacgtc	gccgtccagc	tcgaccagg	1500
tggcaccac	cccggtgaac	agctcc	ccttgc	catggggcg	ttagcttgc	1560
ttgacagtgg	ctgggggtt	cgccgcggg	tttata	agccacagcg	gccactcg	1620
ccataaaaagg	caactttagg	aacggccggg	ggtgatttgc	ttcgagtc	ttattcacc	1680
gccttgcgc	acagtgcagc	attttttac	ccctctccc	ctcctttgc	gggggaaaaa	1740
aaaaaaaaaa	aaaaaaaaaaag	gagaagagaa	aaaaagcgag	cgagagagaa	agcgagattt	1800
aggaagagga	tgaagagttt	ggcgatgggt	gctgggtcc	taggcccaga	tggacaagaa	1860
tagccccgc	ccttgcggac	agtatccat	tcagtgactc	agatcgtca	agcggccgc	1920
tctaacctag	gagcccacca	cagacactat	actataacca	cactgaattt	tctctcctat	1980
ctcaaataat	cctatttctg	ttcatgcctc	agtgcccttat	aacaagcaga	tttctcg	2040
agacatgg	gcttgcgc	gtaatcccag	cacttggc	aagtccggca	gctca	2100
ggtcaggagt	ttgaattc	cctgtccat	atgggtgaaac	cccatctc	ctaaaaatac	2160
aaaaattagc	tggcgtatgt	ggtgtgtgc	tgtactccca	gctactcagg	agactgaggc	2220
aggagaatttgc	tttgcactcg	ggaagcagag	gttgc	gcaagatca	tgcatt	2280
ctccagcctg	ggcaacagag	cgagactccg	tctcaaa	aacaaaacaa	aaacaaaaaa	2340
acaagcagat	ctctcaactt	ggaatgtgc	tgcttctca	tgccaactgg	aaaacttctt	2400
tttcttttag	actcaactca	aatatgatcc	tccctga	ctctccatct	ttcctagaca	2460
caattagcaa	ccagttc	tgtgttcaaa	cagcactcg	tgtgtatgt	tatgtatgt	2520
tatttatcat	tattttatc	accttgc	actctcc	aaggtaatag	ccaggcactt	2580
agtatagagt	aggtatcaac	aattgc	tgaaggata	aatgggtgaa	tgacaactt	2640
gcaaagg	tttcaat	ctattgattt	gaaagaatta	ttatacatta	ttttaat	2700
cotttctagt	tctaataatct	tatgtctata	gtagttcgac	agttacttta	aatcagat	2760
taaatgattt	tcattttaca	aagaacagac	ctgc	caatgtcaca	agactattgt	2820
agagaatagc	agtgaat	gtacatgt	aaatacttgg	caaagtaata	gtgaaaatag	2880
agctgatctc	acaggaatgt	tttggaaattt	aaaacaccgt	atgtaaaata	gtaatctaa	2940
tctgtgggc	catat	ctctcaataa	atattaactt	tgattattt	aaaaatacc	3000
ggtagctt	ttgtgaattt	tacaagaacc	tgtacatgt	tggttgc	ttccgtcaaa	3060
atgagaaaat	tttcttctgc	actgaatttct	tcttttaact	aaattttaag	aaaaaaacca	3120
atgattaaaa	agttagataag	aaaatgataa	tataaaacca	ttagactaaa	aaaaaactag	3180
gttttagatga	taaaaatttgc	aaacccatct	tactatttct	cagatcttgc	gagggcaatt	3240
accattatca	agatagg	gagacttcat	actcttgc	tgtttatgt	tgctgagat	3300
ttaaaagtct	aacttgcatt	ctcagctca	atcatttca	agtaacat	tcacatataa	3360
attatacc	caaggcataa	aaaactccag	aaaatattaa	ctttaaacc	acatattttc	3420
agatattata	catacattt	tattttat	ttcttgaggc	atcctctgtt	tgttatat	3480
atccataataa	aatatgt	gagtcttgc	atctgggg	agttc	tctctacac	3540
tggcaagaaa	ataaaaagg	ttagggaga	tgtgtataaa	gccttcttc	tatgtgtca	3600
ttcatctgg	gggaccta	agcagaat	ataaaaaaaa	aaaggcataa	atcagcagg	3660
aactatgagt	tctaagagag	tgtataattt	cttctctgg	gtcctcaggc	aggccat	3720
acatttca	gaatttcc	atctataaaa	tatgataaa	aaaatattcc	acttgcgtat	3780
ttcttaggaa	taaatggtag	tctggat	aagactatttgc	agcccattca	tattgaaggt	3840
cactgaaagt	tcaagctt	attataacta	atgaaagg	ttcataat	atgactacta	3900
tgaataactat	attcatatac	tattctacat	gaattccacc	tggtaagtc	atttgagat	3960
attcatgtgg	aggagggato	aacaaactgt	ggccat	ccaaatccat	tggat	4020

aaagtttat	tggaacacag	tcacactcat	ttgttatgt	attgtctctg	ctactttgc	4080
actaaaatgg	cagagtttag	tagctgtgt	agataccatg	tggccaaagc	cctttaaaga	4140
aaaaaaagt	gtcaaccctt	aaaataagaga	gtcaacaatc	tggacaattc	agaccccaa	4200
catctcttc	tccctctaca	gcatctttg	aagagtgtga	ataattatgt	caactctttg	4260
ttttcttata	ccaaagtgtt	actgatgaag	cactatggca	aaataatgt	ttcatataaa	4320
atacttttc	ttaaataaat	acaaaataac	gtatttgtaa	aactaaatgt	ggataaaagt	4380
ttctttgtca	ataatgtata	aaatattgag	aagggtgtgg	tttggacaat	catagcttac	4440
tggttttca	gctaattggca	aaataaattt	ccatcaaatg	tcctgtcat	tatcacaatg	4500
tgtgtgtcta	aacaccttgg	atgctgtcc	cctattttgt	cctaagtgc	tcaatggatt	4560
tcttgagaaa	tttgcatacc	cagtggtag	gacttttctc	tattcatcca	catacggtga	4620
ttcacccctgt	tttactcatt	tataaggatt	aatccacagc	ctctatttgag	accatgttct	4680
tcccaggttt	agactgtctt	ctaaaaaga	ggaaaaaaat	ggtcatcaat	actcagttgg	4740
aaaaaaatcca	tttatacac	caaatcattt	ctgctggca	acactaggt	gtaggactgg	4800
aatggagtca	gagagaagga	aaacttcac	tttttagttt	atgcattgt	actatttggaa	4860
tgttacatc	aaagcatgt	tttaggaatcc	attttaatga	acaattggac	tagaaattca	4920
gctgtgcctg	tggcttaat	ttgaaagcct	cctggccagc	tcactatgt	cctagaaaact	4980
cagttcctgc	cagtcttggc	agacaccatt	gcatattattt	accctgcatt	actttgttg	5040
cgagggtact	tgtcgcaagt	cacgacettc	actgtcttct	ctagacggcg	atacagtcaa	5100
ctgtcttga	cctttgttac	tactctttc	cgatgtatgat	gtcgcaactt	ttctatgctg	5160
tctcaatgtt	agaggcatat	cagtctccac	tgaagccat	ctatctgtga	cggcatctt	5220
attcacatta	tcttgtacaa	ataatcctgt	taacaatgt	tttatatcct	gtaaagaatc	5280
catttcaaa	atcatgtcaa	ggtcttctcg	agaaaaaaatc	agtagaaaata	gctgttccag	5340
tctttcttagc	cttgatttca	cttctgtcag	atgtgcccta	gtcagcggag	accttttgg	5400
tttgggagag	tagcgacact	cccagttgtt	cttcagacac	ttggcgcact	tcggtttttc	5460
tttgggagcac	ttgagctttt	taagtccgca	aatatcgat	gcttgttgcg	tagaaagacag	5520
tagcttcatg	gtggcggatc	ccgataagcc	agtaaggcgt	gggttctcta	gttagccaga	5580
gagctctgct	tatatagacc	tcccacgta	cacgcctacc	gcccatttgc	gtcaatgggg	5640
cggagttgtt	acgacatttt	ggaaagtccc	gttgattttg	gtgccaaaac	aaactcccat	5700
tgacgtcaat	gggggtggaga	cttggaaatc	ccccgtggatc	aaaccgttat	ccacgccccat	5760
ttgatgtactg	ccaaaaccgc	atcaccatgg	taatagcgat	gactaatacg	tagatgtact	5820
gccaagtagg	aaagtcccat	aaggtcatgt	actgggcata	atgccaggcg	ggccatttac	5880
cgtcatttgc	gtcaataggg	ggcgtacttg	gcatatgtata	cacttgatgt	actgccaagt	5940
gggcaggta	cgtaaataac	tccacccatt	gacgtcaatg	gaaagttccct	attggcgat	6000
ctatggAAC	atacgtcatt	attgacgtca	atgggcgggg	gtcggtggc	ggtcagccag	6060
gcggggccatt	taccgttaat	tatgtaacgc	ggaactccat	atatggct	tgaactaatg	6120
accccgtaat	tgattactat	taataactag	tcaataatca	atgtcaacgc	gtatatctgg	6180
cccgtaatc	gogaaggcgc	gcaaaacgcc	taaccctacc	cgggctgcag	gaattctacc	6240
gggttagggga	ggcgctttc	ccaaggcagt	ctggagcatg	cgctttagca	gccccgctgg	6300
cacttggcgc	atcacaagt	gcctctggcc	tcgcacacat	tccacatcca	ccggtagcgc	6360
caaccggctc	ccttctttgg	tggcccttc	gcccacacat	ctactcctcc	cctagtgcagg	6420
aagttcccc	ccgccccgca	gctcgctcg	tgcaaggacgt	gacaaatgg	agtgcacgt	6480
ctcaactagtc	tcgtcgacat	ggacaaggac	cgttgaccaa	tggaaagcggg	taggccttt	6540
gggcagcggc	caatagcgc	ttggctccct	cgctttctgg	gctcagaggc	tgggaagggg	6600
tgggtccggg	ggcgggctca	ggggcgggct	cagggggcggg	gccccggcga	aggtccctccg	6660
gaccggcat	tctgcacgct	tcaaaacgca	acgtctggcg	cgctgttctc	ctttcctca	6720
tctccgggccc	tgcacactgc	atgaaaaagc	ctgaactcac	cgcgacgtct	gtcgagaagt	6780
ttctgtatcg	aaagttcgac	agcgtctccg	acctgatgca	gctctcgag	ggcgaagaat	6840
ctcggtctt	cagttcgat	gtaggaggc	gtggatatgt	cctgcgggt	aatagctgcg	6900
ccgatggttt	ctacaaagat	cgttatgttt	atccggactt	tgcacatggcc	gctgtcccg	6960
ttccggaaatg	gttgcacatt	ggggaaatca	gctggacgtt	gacatttgc	atctcccgcc	7020
gtgcacagg	tgtcacgtt	caagacactgc	ctgaaacccga	actgcccgt	gttctgcagc	7080
cggtcgcgg	ggccatggat	gcatcgctg	cgccgcgtat	tagccagacg	agcgggttcg	7140
gcccatccgg	accgcaagga	atcggtcaat	acactacatg	gcgtgattt	atatgcgcga	7200
ttgctgatcc	ccatgtgtat	caactggcaaa	ctgtgtatgg	cgacaccgtc	agtgcgtccg	7260
tgcgcaggc	tctcgatgag	ctgatgttt	ggcccgagga	ctgccccgaa	gtccggcacc	7320
tctgtcacgc	ggatttcggc	tccaacaatg	tcctgacgg	caatggccgc	ataacagcgg	7380
tcattgtactg	gagcgaggcg	atgttcgggg	attcccaata	cgaggtcgcc	aacatcttct	7440
tctggaggcc	gtgggtggct	tgtatggagc	agcagacgcgc	ctacttcgag	cgaggccatc	7500

cggagcttgc	aggatcgccg	cggctccggg	cgtatatgct	ccgcattgg	tttgaccaac	7560	
tctatcagag	cttggttgac	ggcaatttcg	atgatgcagc	ttgggcgcag	ggtcgatg	7620	
acgcaatcgt	ccgatccgg	gccgggactg	tcgggcgtac	acaatcgcc	cgcagaagcg	7680	
cgccgcgtc	gaccgatgg	tgtgtagaag	tactcgccg	tagtgaaac	cgacgcccc	7740	
gcactcgcc	gagggcaaag	gaatagaga	tcagaaaatt	gatgatctat	taaacaataa	7800	
agatgtccac	taaaatggaa	gttttcttgc	tcataacttt	ttaagaagg	tgagaacaga	7860	
gtacctacat	tttgaatgg	aggattgg	ctacgggg	gggggtgggg	ttggattaga	7920	
taaatggct	ctctttact	aaggcttt	actattgct	tatgataatg	tttcatagt	7980	
ggatatcata	atttaaaca	gaaaaaccaa	attaaggcc	agctcatcc	tcccactcat	8040	
gatctataga	tctatagatc	tctcggtgg	tcatttttt	tcttgcatt	cccactttgt	8100	
ggttctaagt	actgtgg	ccaaatgt	cagttcata	gcctgaagaa	cgagatcagc	8160	
agcctctgtt	ccacatacac	ttcatttca	gtattttt	gccaagttt	aattccatca	8220	
gattttagt	gacatctt	ttgttata	gatcatca	ttctgcata	taagcttac	8280	
gataccgtcg	acctcgagg	ggggcccg	accagctt	tgttccctt	agtgggg	8340	
aattgcgcgc	ttggcgtat	catggcata	gctgttcc	gtgtaaaatt	gttatccgc	8400	
cacaattcca	cacaacatac	gagccgga	cataaaagt	aaagcctgg	gtgccta	8460	
agttagctaa	ctcacattaa	ttgcgttgc	ctca	gctttccagt	cgggaaac	8520	
gtcgtgccag	ctgcattaa	aatcgcc	acgcgcggg	agaggcg	tgcgtattgg	8580	
gctctttcc	gttcctcgc	tcactgact	gtcgctcg	gtcg	tcggcgagc	8640	
ggtatcagct	cactcaaagg	cgtaatac	gtatccaca	aatcagg	ataacgcagg	8700	
aaagaacatg	tgagcaaaag	gccagcaaa	ggcaggaa	cgtaaaaagg	ccgcgttgc	8760	
ggcgttttc	cataggctcc	gccccctga	cgagcatc	aaaaatcgac	gctcaagtca	8820	
gaggtggcga	aacccgacag	gactataaag	ataccaggcg	tttccccctg	gaagctccct	8880	
cgtgcgtct	cctgttccg	ccctgcgc	taccggata	ctgtccgc	tttcccttc	8940	
gggaagcgtg	gcttttctc	atagctc	ctgtagg	ctcagg	tgtaggtcg	9000	
tcgctccaag	ctgggctgt	tgac	ccccgtt	ccgcacc	gcgcctt	9060	
cggtaact	cgttttgc	ccaaacccgt	aa	acacacgac	ttatcgcc	9120	
cactggtaac	aggattagc	gagcggat	tgtaggcg	gtc	acagat	9180	
gtggcctaac	tacggctaca	ctagaagg	at	tcgtcg	tgcgtaa	9240	
agttac	ggaaaaag	ttgttagt	ttgtatcc	aa	acaaacca	ccgctgg	9300
cgggtttt	tttgttgc	agcagat	tacgcgc	aaaaaaggat	ctcaaga	9360	
tccttgc	tttctacgg	ggtctgac	tcagtg	aaaaactc	gttaagg	9420	
tttgttgc	agattatcaa	aaaggat	cac	ttttaaatt	aaaaat	9480	
ttttaatca	atctaaag	at	atgat	aa	cttgc	9540	
cagtgg	cctatctc	cgatct	tttgc	ttccat	tttgc	9600	
cgctgttgc	ataactac	ta	ccat	ggcccc	ctgca	9660	
accgcgagac	ccacg	cg	ttt	at	aaacc	9720	
ggccgagcgc	agaagtgg	ctgca	at	cc	cagcc	9780	
ccggaaagct	agaga	gtcgcc	at	tt	tttgc	9840	
tacaggc	gtgg	ctac	cg	cc	at	9900	
acgatcaagg	cgagt	at	tc	cc	cc	9960	
tcctccgatc	gttgc	gt	tc	cc	cc	10020	
actgcataat	tctttact	tc	at	cc	cc	10080	
ctcaaccaag	tcattct	at	ttt	cc	cc	10140	
aatacggat	aataccgc	ca	ttt	cc	cc	10200	
ttcttcggg	cgaaaact	ca	ttt	cc	cc	10260	
cactcg	cccaact	cc	ttt	cc	cc	10320	
aaaaacagga	aggcaaaat	cc	ttt	cc	cc	10380	
actcatactc	ttctttt	aa	ttt	cc	cc	10440	
cggatacata	tttgaatgt	ttt	ttt	cc	cc	10500	
ccgaaaat	ccac	ttt	ttt	cc	cc	10514	

<210> 51
<211> 10247
<212> DNA
<213> Artificial Sequence

ggaaacctgt	cgtgccagct	gcattaatga	atcgcccaac	gcgcggggag	aggcggtttg	2940
cgtattggc	gctttccgc	ttcctcgctc	actgactcgc	tgcgctcggt	cgttcggtctg	3000
cggcgacgg	tatcagctca	ctcaaaggcg	gtataacaggt	tatccacaga	atcagggat	3060
aacgcaggaa	agaacatgtg	agcaaaaaggc	cagcaaaaagg	ccaggaaccg	taaaaaggcc	3120
gcgttgcgtgg	cgttttcca	taggctccgc	ccccctgacg	agcatcacaa	aatcgacgc	3180
tcaagtccaga	gttggcgaaa	cccgcacagga	ctataaagat	accaggcggtt	tccccctgga	3240
agctccctcg	tgcgctctcc	tgttccgacc	ctgcccgtta	ccggataccct	gtccgccttt	3300
ctccccctcg	gaagcgtggc	gcttttcata	agtcacgcgt	taggttatct	cagttcggtg	3360
taggtcgttc	gctccaaatcg	gggtctgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	3420
gccttatccg	gtaaatctcg	tctttagtcc	aaccccggtaa	gacacgactt	atcgccactg	3480
gcagcagcca	ctggtaacag	gattagcaga	gcgaggatatg	taggcgtgc	tacagagttc	3540
ttgaagtgg	ggcctaacta	cggctacact	agaaggacag	tattttgtat	ctgcgctctg	3600
ctgaagccag	ttaccccg	aaaaagagtt	ggtagcttct	gatccggcaa	acaaaccacc	3660
gctggtagcg	gtggttttt	tgttgcaag	cagcagattt	cgccgagaaa	aaaaggatct	3720
caagaagatc	cttgatctt	ttctacgggg	tctgacgctc	agtggAACGA	aaactcacgt	3780
taagggattt	tggtcatgag	attatcaaaa	aggatctca	cctagatct	tttaaattaa	3840
aaatgaagtt	ttaaatcaat	ctaaagtata	tatgagtaaa	cttggctgt	cagttaccaa	3900
tgcttaatca	gtgagggcacc	tatctcagcg	atctgtctat	ttcggttcatc	catagttgcc	3960
tgactccccg	tctgttagat	aactacgata	cgggaggggct	taccatctgg	ccccagtgct	4020
gcaatgatac	cgcgagaccc	acgctcaccg	gctccagatt	tatcagcaat	aaaccagcca	4080
gccggaaggg	ccgagcgcag	aagtggctt	gcaactttat	ccgcctccat	ccagtctatt	4140
aattgttgcc	gggaagctag	agtaagtatg	tcgcccgtta	atagtttgcg	caacgttgg	4200
gccattgcta	caggcatcgt	ggtgtcacgc	tcgtcggtt	gtatggctt	attcagctcc	4260
ggttcccaac	gtcaaggcg	agttacatga	tcccccatgt	tgtgcaaaaa	agcggttagc	4320
tccttcggc	ctccgatcgt	tgtcagaagt	aagttggccg	cagtgttata	actcatgg	4380
atggcagcac	tgcataattc	tcttactgtc	atgcccattcg	taagatgttt	ttctgtgact	4440
ggtgagact	caaccaatgc	attctgagaa	tagtgtatgc	ggcgaccgag	ttgtcttgc	4500
ccggcgccaa	tacggataa	taccgcgc	catagcagaa	ctttaaaatg	gtctcatcatt	4560
gaaaaacgtt	cttcggggcg	aaaactctca	agatcttac	cgctgttgag	atccagttcg	4620
atgttaacca	ctcggtcacc	caactgtatc	tcagcatctt	ttactttcac	cagcgtttct	4680
gggtgagcaa	aaacagggaa	gcaaaatggc	gcaaaaaaagg	gaataagggc	gacacggaaa	4740
tgttgaatac	tcatactctt	ccttttcaa	tattattgaa	gcattttatca	gggttattgt	4800
ctcatgagcg	gatacatatt	tgaatgtatt	tagaaaaata	aacaaatagg	ggttccgcgc	4860
acatcccccc	gaaaagtgc	acctaattt	taagcgtaa	tatttgtta	aaattcgct	4920
taaattttt	ttaaatcgc	tcattttta	accaataggc	cgaaatcg	aaaatccctt	4980
ataaaatcaaa	agaatagacc	gagatagggt	tgagtgtt	tccagttgg	aacaagagtc	5040
cactattaaa	gaacgtggac	tccaacgtca	aaggcgaaaa	aaccgtctat	caggcgatg	5100
gcccactacg	tgaaccatca	ccctaataa	gtttttggg	gtcgagggtc	cgtaaagcac	5160
taaattcgaa	ccctaaagggg	agcccccgat	ttagagctt	acggggaaag	ccggcgaaacg	5220
tggcgagaaa	ggaagggaaag	aaagcgaaag	gagcggggc	tagggcgct	gcaagtgtag	5280
cggtcacgct	gcbcgttaacc	accacacccg	ccgcgttta	tgcggcgt	caggcgctg	5340
cccatcgcc	attcaggctg	cgcaactgtt	ggaaaggggcg	atcggtcg	gcctcttcgc	5400
tattacgcca	gctggcgaaa	ggggatgtg	ctgcaaggcg	attaagtgg	gtaacgccc	5460
ggtttccca	gtcacgacgt	tgtaaaacga	cgcccaagt	gcccgcgtaa	tacgactcac	5520
tatagggcga	atggagctc	caccgcgtg	cgccggggcc	atgcaggcca	cgacatgata	5580
agatacattt	atgagtttgg	acaaaccaca	actagaatgc	agtggaaaaaa	atgttttatt	5640
tgtgaatattt	gtgtatgtat	tgcttttattt	gtaaccat	taagatgc	taaacaagtt	5700
cttgcatacg	tgcgtccatgc	cgagagtat	cccgccggcg	gtcacgaact	ccagcaggac	5760
catgtatcg	cgttttcgt	tgggttctt	gtcaggggcg	gactgttag	tcaagttgt	5820
gttgcggcc	agcagcacgg	ggccgtcgcc	gatgggggtt	ttctgtgtt	agtggtcgg	5880
gagctgcacg	ctgcgtctt	cgatgtgt	gcccgtt	aagttcacct	tgatggcg	5940
cttctgtt	tccggccat	tatagacgtt	gttgcgtt	tagtttact	ccagcttgc	6000
ccccagatg	ttggcgctt	ccttgaagtc	gatggcccttc	agctcgatgc	ggttaccac	6060
gggtgcggcc	tgcgacttca	cctcgccgc	ggtcttgc	ttgcgtcg	ccttgaagaa	6120
gatggtgcc	tcctggacgt	agccttcgg	catggcgac	ttgaagaagt	cgtgtcg	6180
catgtggctcg	gggttagcg	cgaagcaact	caggccgt	ccgaagggtt	tcacgagggt	6240
ggggccaggc	acggggcagct	tgccgtt	gcagatgaac	ttcagggtca	gcttgcgc	6300
ggtggcatcg	ccctcgccct	cgccggacac	gctgaactt	tggccgttta	cgtcggcg	6360

cagctcgacc	aggatggca	ccaccccggt	gaacagctcc	tcgcccgtgc	tcaccatggt	6420
ggcgtagct	tgatttgaca	gtggctgggg	gttgcgcgc	cgggttttat	aggaagccac	6480
agcggccact	cgagccataa	aaggcaactt	taggaacggc	gggggggtat	tggattcgag	6540
tcgttattc	acccggcctt	ccgcacagt	cagcattttt	ttacccctc	ccccctcctt	6600
ttgcggggga	aaaaaaaaaa	aaaaaaaaaa	aaaggagaag	agaaaaaaag	cgagcgagag	6660
agaaaaggcg	attgaggaag	aggatgaaga	gttggcgat	gggtgtcggt	tccgtaggcc	6720
cagatggaca	agaatagccc	ccgccttgc	ggacagttt	ccattcagt	actcagatca	6780
gatcaagcgg	ccggcagtgt	gatggatac	tcgagaattt	cagcacact	gccccgtta	6840
ctagtggatc	ctcaatcgt	ccatcacct	ggtaactttt	ctacccaat	ctgaaagtga	6900
atatttggaa	aactatgttt	ctaactggaa	gaattctgaa	actatgttct	tctacaacat	6960
gtagacatg	tgtcacaatg	aaaatgaaag	ttccttctga	agagctgtaa	ccctacatag	7020
cctaggatgc	catttatact	ttctctagcc	cagaggctt	gtatttctgc	tctcaggcata	7080
cattactaca	taaccatcaa	accaactgcc	cggaaatttgc	gagcaatttga	tcataacata	7140
cagacctgt	tcagatataat	aaattacatt	tttatttttta	tcgtacccaa	acatacagg	7200
gaaaacttcc	aaattcaggc	aactgacat	ccaggtgcct	acctatctgt	gaagaccaga	7260
gcactgtaa	acacggata	gctgggtgaa	aagtggaaaa	tccagtgtgt	tagttgttt	7320
gattatacag	aatcatagaa	tttcaaaatgt	gaaaggacat	ggtgatca	gatttcagt	7380
actatccaa	aacatttagca	agagccctta	gggttcccta	ccagttcctt	cagaagccctc	7440
cttcaaaagct	gtgagtgcat	ggccaaaggg	gaagctacat	tcatcaacca	gactaatcag	7500
agaccaggga	gcagaaataa	ggtaaacacg	gaatggagta	catttgc	gggggacgtt	7560
catgtttct	tgcgatact	gaaaacagct	gggtgtgtat	tttggtataa	cactcagagc	7620
ctctttgtat	aaaaaggcat	tctatatgaa	tggaaaccct	gttggcttca	catttgtcaa	7680
agcaatggat	tttaaatagt	gactcaactt	ccttagaaaac	tttgcatttt	gaaatgtttt	7740
cttcatacag	ctaacaggca	aaactaatca	aaaccttgc	gttagtctga	tgtttatttt	7800
ctgtgcagca	aacaaactcc	cataaaacc	tctgagggtt	tcttagttat	gtttgcatt	7860
ttttaagtgt	tgaactgcct	aagtctctga	tgtcatgaca	cttaatttcat	atattttctt	7920
gcttttgc	gaactgaaca	aacactccct	atcaagattt	accctctta	gaaataaaat	7980
aaatttgc	atgaagtcaa	aatatgacat	atcaagttt	tgatagctt	cttgcatt	8040
tatggggaga	gaaaatttata	ccatccatac	taataggtt	ggggttttaa	taatatttgc	8100
ctataaaca	taagcaattt	cagatttatt	ttaagtttac	gtatgcatt	cctgttaagca	8160
acaggaaata	tattttacat	caatgtatgc	actgccttgc	agtaacacg	attatatttt	8220
acatgaatgt	gtgaagtgtat	attttttaa	taggcaaaaa	ttagttttaa	aagcgaaact	8280
ggcatgc	cctaggctgt	ctcagtgata	atgacccat	agtcaatgt	ttcttgcatt	8340
agatgaaata	agtttattgc	aacttgattt	catcacctaa	aatacaaccc	ttgacttgg	8400
agcttgc	tatgttttat	gaggtgtat	ataaaaagg	aactgtat	actgaaccag	8460
caaattattt	tgcataatccat	gactacat	gtatacagca	gatgtat	attgttgg	8520
aaagttggac	caaatttca	tcctgttgc	tactgtgtat	agctcttcc	aaaatacatt	8580
cttgccttta	gaggaactga	cccattat	atcatacaga	tagaattttt	tgaacttatt	8640
ttaaagggtt	ggttattat	caatgtat	ctctcaat	tcttgcatt	ttccactgt	8700
aaaaccaat	atctgcatt	taggcaaaa	attaatattt	ctcagatgt	ttagagtaca	8760
tatagatgaa	gactgcctt	ttacttagat	agttttgaga	gagatggaaa	caaaagaaag	8820
aaaaaaatagg	atagctgtat	aaatgaaact	tcaaattttaa	tttgatcat	ttggcatttt	8880
ttattcc	agtgcatt	aaataactga	ttaatactat	acacatataa	tactatcaag	8940
ttacacc	aaatttattt	gtcagccagg	caaccaatca	gaaccaata	gtgaccaaga	9000
attattttt	caaacacgtt	tgtcttcc	gaccctact	agttgttcc	acagatcat	9060
aaaacatgg	tctttctca	ggtcttctac	atcatttcc	ttaagtttgc	acccaaagtc	9120
ataagcttga	ctttagacac	aggtttcaaa	agtagaaat	tttatttct	tatgtttct	9180
tggcttggaa	gttttgc	actataaaaa	taatttgc	tacccttcc	acaactcat	9240
gacgttcc	gaatgttgc	aaacgaaat	tgcgtttgc	ataataaaca	gaaaaaaaaa	9300
aaaaaaaaaa	aaaattaaat	gaatctttaat	tctcatat	catttttttt	ttcagaccc	9360
gacatcttca	ggatttctt	acattttaag	cagaatgaca	ctactgtat	acaaaaat	9420
attttaattt	tgtttttat	gcatgaggca	atataaaaa	taatcatat	gatcattata	9480
tttgcattt	caactaaat	tttccat	tcccccagat	tttcaatgc	tctttctcc	9540
tcatggcaga	tgcccaacca	cagagagaaa	ggttctataa	taaaacttta	aaaacactga	9600
cctaaaagag	aatatgtttt	tcgcccatt	atattctaa	ttctggcatt	ttgtggaaaa	9660
actgtacaca	tttttggca	tggctttagt	gtccactgtt	ctggcagct	taatgagaat	9720
aataaagtat	atgggtat	ctgtttactt	accattat	gacatcccc	ctgcggta	9780
cggcgatata	gtcaactgtc	tttgacctt	gttactact	tcttccat	atgtatgtc	9840

acttattctca	tgttgtctca	atgttagagg	cataatcagtc	tccactgaag	ccaatctatc	9900
tgtgacggca	tctttattca	cattatcttg	tacaataaat	cctgttaaca	atgcctttat	9960
atcctgtaaa	gaatccattt	tcaaaatcat	gtcaagggtct	tctcgaggaa	aaatcagtag	10020
aaatagctgt	tccagtcttt	ctagccttga	ttccacttct	gtcagatgtg	ccctagtcag	10080
cgagacacctt	ttggtttgg	gagagtagcg	acactcccag	ttgttcttca	gacacttggc	10140
gcacttcggt	ttttcttgg	agcacttgag	cttttaagt	cggcaaataat	cgcacatgttg	10200
ttcgatagaa	gacagtagct	tcatggtggc	aagcttaagt	ttaaacacg		10247

<210> 52
<211> 10514
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant DNA vector sequence

<400> 52						
ctaaattgtta	agcgttaata	ttttgttaaa	attcgcttta	aattttgtt	aaatcagctc	60
attttttaac	caataggccg	aaatcgccaa	aatcccttat	aaatcaaaag	aatagaccga	120
gatagggttg	agtgttgttc	cagtttggaa	caagagtcca	ctattaaaga	acgtggactc	180
caacgtcaaa	gggcgaaaaaa	ccgtctatca	gggcgatggc	ccactacgtg	aaccatcacc	240
ctaatacgat	tttttggggt	cgaggtggcg	taaagcacta	aatcggaaacc	ctaaagggag	300
cccccgattt	agagcttgcac	ggggaaagcc	ggcgaacgtg	gcgagaaaagg	aagggaaagaa	360
agcgaaaagga	gccccgccta	gggcgcgtggc	aagtgttagcg	gtcacgctgc	gcgttaaccac	420
cacaccggcc	gctctaatg	cgccgtacata	gggcgcgttcc	cattcgcctt	tcaggctgcg	480
caactgttgg	gaaggcgat	cggtgcgggc	cttttcgtctt	ttacgcgcag	ttgcgaaagg	540
gggatgtgct	gcaaggcgat	taagttgggt	aacgccagg	ttttcccagt	cacgacgttg	600
taaaaacacg	gccagtggac	gcgcgtataa	cgactcacta	tagggcgaat	tggagctcca	660
ccgcggcgcg	gccggggccat	gcaggccaga	catgataaga	tacattgatg	agtttggaca	720
aaccacaact	agaatgcagt	aaaaaaatg	ctttatttg	gaaatttgtg	atgctattgc	780
tttatttgta	accattataa	gctgcaataa	acaagttgtc	gagttacttg	tacagctcg	840
ccatgcgcag	agtgtatcccg	gccccgcgtca	cgaactccag	caggaccatg	tgatcgcgct	900
tctcggttgg	gtctttgtc	agggcgact	ggtagctcg	gtagtggtt	tcgggcagca	960
gcacggggcc	gtccgcgtat	gggggtttct	gctggtagtg	gtcggcgagc	tgcacgctgc	1020
cgtcctcgat	gttggcggg	atcttgaagt	tcaccttgc	gccgttctt	tgcttgcgg	1080
ccatgatata	gacgttgcgg	ctgttgcgt	tgtactccag	cttgcgtcccc	aggatgttgc	1140
cgtcctcctt	gaagtgcgt	cccttcagct	cgatgcgg	caccagggt	tcgcctcgt	1200
acttcacctc	ggcgcgggtc	ttgttagttgc	cgtcgtcctt	gaagaagatg	gtgcgtcct	1260
ggacgttagcc	ttcgggcgt	gcggacttga	agaagtcgt	ctgcttcatg	ttgtcggggt	1320
agcgggcgaa	gcactgcagg	ccgtaggccg	agtggtcac	gagggtggc	cagggcacgg	1380
gcagctgtcc	gggtgtgcag	atgaacttca	ggtgcagctt	gccgttagtg	gcacgtccct	1440
cgccctcgcc	ggacacgctg	aacttgcggc	cgttacgtc	gccgtccagc	tcgaccagga	1500
tgggcaccac	cccggtgaac	agctccctcg	ccttgcac	catggtggcg	ttagcttgc	1560
ttgacagtgg	ctgggggttgc	cgccgcggg	ttttatagga	agccacagcg	gccactcgag	1620
ccataaaaagg	caactttagg	aacgggggg	ggtgatttgc	ttcgagtcgt	ttatttcaccg	1680
gccttgcgc	acagtgcagg	attttttac	ccctctcc	ctccttcgtc	ggggaaaaaa	1740
aaaaaaaaaa	aaaaaaaaaa	gagaagagaa	aaaaagcgag	cgagagagaa	agcgagattg	1800
aggaagagga	tgaagagttt	ggcgtgggt	gtctgggtcc	taggcccaga	tggacaagaa	1860
tagccccgc	ccttgcggac	agtatccat	tcagtgcgtc	agatcgtca	agcggccgcg	1920
tctaacctag	gagcccacca	cagacactat	actataacca	cactgaattt	tctctcctat	1980
ctcaaataat	cctattcttgc	ttcatgcctc	agtcgttat	aacaagcaga	tttctcggcc	2040
agacatggtg	gcttgcgcct	gtaatcccag	cacttggcc	aagtccggca	gctcaacttgc	2100
ggtcaggagt	ttgaatttgc	cctgtccaaat	atggtgaaac	cccatctcta	ctaaaaataac	2160
aaaaatttagc	ttggcatgg	gggtgtgtcc	tgttagtccca	gctactcagg	agactgaggc	2220
aggagaatttgc	tttgcactcg	ggaagcagag	gttgcagtgc	gccaagatca	tgccatttgc	2280
ctccagccgt	ggcaacagag	cgagactccg	tctcaaaacaa	aacaaaacaa	aaacaaaaaa	2340

acaaggcagat ctctcaactt ggaatgctgc tgcttctcca tgcccaactgg aaaacttctt	2400
tttcttttag actcaactca aatatgatcc tcacctgaagt ctctccatct ttccttagaca	2460
caattagcaa ccagttcctc tgtgttcaaa cagcacttcg tgtgtatgtg tatgtatgt	2520
tatttatcat tatttatac acctttgatc actctccctg aaggtaatag ccagggactt	2580
agtatagagt aggtatcaac aattgtttat tgaaggaata aatgggtgaa tgacaacttt	2640
gcaaaggttc aatttttaat ctattgatt gaaagaata ttatacatta ttttaaatat	2700
cctttctagt tctaatactt tatgtctata gtatgtcgac agttactta aatcagat	2760
taaatgattt tcaatttaca aagaacagac ctgccttcttcaatgtcaca agactattgt	2820
agagaatagc agtggaaataa gtacatgtga aataacttgg caaagaataa gtgaaaatag	2880
agctgatctc acaggaatgt tttggaaaattt aaaaacaccgt atgtaaaaataa gttaatctaa	2940
tctgtggca catatttaaa ctctcaataa atattaactt tgattattta aaaaataacct	3000
ggtagcttt ttgtgaattt tacaagaacc tgatcatgat tggttgcatt ttccgtcaaa	3060
atgagaaaaat tttcttctgc actgaattct tcttttaact aaattttaa aaaaaaacca	3120
atgattaaaa agtagataag aaaaatgataa tataaaacca ttagactaaa gaaaaactag	3180
gtttagatga taaaatttga aaaccatct tactatttct cagatcttg gagggcaatt	3240
accattatca agataggtgg gagacttcat actcttgaat tgtttaatgt tgctgagagt	3300
ttaaaagtct aacttgaatt ctcagctcta atcatttcaa agtaacatatac tcaatataa	3360
attatacctc caaggcaata aaaaactccag aaaaatattaa cttaaacc acatattttc	3420
agatattata catacaattt tatttggat ttcctgaggc atcctctgtt tggttatatga	3480
atccataataa aaatgttagaa gagtctgtt atctgggtgt agttcagttt ttcctcacac	3540
tggcaagaaa ataaaagggtg ttagggaaaga tggttataaa gccttcttc tataatgtgca	3600
ttcatctgga gggacctaattt agcagaaatc ataaaaaaaaa aaaggcataa atacagcagg	3660
aactatgagt tctaagagag tgtataattt ctttctctgg gtcctcaggc aggccatatc	3720
acatttcaact gaattttctt atctataaaa tatgaataac aaaatattcc acttgtgtat	3780
ttcttaggaa taaatggtag tctggatatg aagactattt agccccattca tattgaaggt	3840
cactgaaagt tcaagcttta attataacta atgaaagggtt ttcataatagt atgactacta	3900
tgaataactat attcatatac tattctacat gaattccacc tggtgaagtc atctgagagt	3960
attcatgtgg aggagggtac aacaaactgt ggcccattgg ccaaattccat tggttttat	4020
aaagtttat tggAACACAG tcacactat ttgtttatgt attgtctctg atcttttgc	4080
actaaaatgg caggttgag tagctgttg agataccatg tgcccaaagc cttttaaaga	4140
aaaaaaatgtt gtcaaccctt aaaaatagaga gtcaacaatc tggacaattc agaccccaaa	4200
catctcttc tccctctaca gcatctttt aagagtgtga ataattatgt caactctttg	4260
tttcttata ccaaagtgtt actgtatggc aactatggca aaataatgt ttcatataaa	4320
atacttttc taaaataat aaaaaatac gtatttgtaa aactaaatgt ggataaaagt	4380
ttctttgtca ataatgtat aataatttgg aagggtgtgg tttggacaat catagcttac	4440
tggttttca gctaatggca aaataaaattt ccatcaatag tcctgtgcat tatacataatg	4500
tgtgtgtctc aacaccttgg atgctgttcc cctattttgt cctaagttgc tcaatggatt	4560
tcttgagaaaa ttgttaaccc cagtggttag gactttctc tattcatcca catacggttgc	4620
ttcacccctgt ttactcatt tataaggatt aatccacacg ctctatttgg accatgttct	4680
tcccaggttt agactgtctt ctcaaaaaaga gggaaaaaat ggtcatcaat actcagttgg	4740
aaaaaaatcca ttttatacac caaatcattt ctgctggca acaactaggtt gttagactgg	4800
aatggagtca gagagaagga aaactttcac tttttatgtt atgcattgtt actatttggaa	4860
tgtttacatc aaagcatgtt ttaggaatcc attttatgtt acaattggac tagaaattca	4920
gctgtgcctg tggctttat ttgaaaggctt cctggccagc tcactatgtt cctagaaact	4980
cagttctcgc cagtcttggc agacaccatt gcatattttt accctgcattt actttgttgc	5040
cgaggctact tgcgtcaagt cacgacccctt actgtcttcc cttagacggcg atacagtcaa	5100
ctgtcttgc cctttgttac tactcttcc cgatgtatgtt gtcgcactt ttctatgtct	5160
tctcaatgtt agaggcatat cagtctccac tgaagccat tcatctgtt gggcatctt	5220
attcacatca tcttgcataa ataatctgt taacaatgtt ttatattctt gtaaaatgtt	5280
cattttcaaa atcatgtcaa ggtcttctcg agggaaaaatc agtagaaataa gctgttccag	5340
tctttcttagc ttgttccat cttctgtcag atgttccctt gtcagcggag accttttgg	5400
tttgggagag tagcgcacact cccagttgtt cttcagacac ttggcgcact tcggttttt	5460
tttggagcac ttgagctttt taagtccggca aatatcgcat gcttgttgc tagaaagacag	5520
tagttcatg gtggcggatc ccgataagcc agtaaggcattt gggtttctca gtttagccaga	5580
gagctctgct tatatagacc tcccacccgtt cacgocctacc gcccatttgc gtcataatgggg	5640
cgaggtgtttt acgacattttt ggaaaggcccc gttgattttgc gtcacaaaac aaactcccat	5700
tgacgtcaat ggggtggaga cttggaaatcc cccgtgagtc aaaccgctat ccacgccccat	5760
tgtatgtactt cccaaaccgc atcaccatgg taatagcgat gactaatacg tagatgtact	5820

gccaagttagg	aaagtcccat	aaggcatgt	actgggcata	atgccaggcg	ggccatttac	5880
cgtcattgac	gtcaataggg	ggcgtacttg	gcatatgata	cacttgatgt	actgccaagt	5940
gggcagttt	ccgtaaatac	tccaccatt	gacgtcaatg	gaaagtccct	attggcgtta	6000
ctatggAAC	atacgtcatt	attgacgtca	atgggggggg	gtcggtggc	ggtcagccag	6060
gcgggcatt	taccgttaatg	tatgtaacgc	gaaactccat	atatggcta	tgaactaatg	6120
accccgtat	tgattactat	taataactag	tcaataatca	atgtcaacgc	gtatatctgg	6180
cccgtaatc	gcaaggcaggc	gcaaaacgcc	taaccctacc	cgggctgcag	gaattctacc	6240
gggttagggg	ggcgctttc	ccaaggcagt	ctggagcatg	cgcttagca	gccccgtgg	6300
cacttggcgc	atcacaagtg	gcctctggcc	tgcacacat	tccacatcca	ccggtagcgc	6360
caaccggc	ccttcttgg	tggcccttc	gcccacett	ctactctcc	cttagtcagg	6420
aagttcccc	cegccccgca	gctcgcgtcg	tgcaggacgt	gacaaatgga	agtgcacgt	6480
ctcaactagtc	tcgtcagat	ggacaagcac	cgttgagcaa	tggaaagcggg	taggccttt	6540
gggcagcggc	caatagcagc	ttggctcctt	cgcttctgg	gctcagaggc	tgggaagggg	6600
tgggtccggg	ggcgggctca	ggggcgggct	caggggcggg	gcccgcgca	aggtcctccg	6660
gacccggcat	tctgcacgct	tcaaaagcgc	acgtctccgc	cgctttctc	ctttcctca	6720
tctccggcc	ttcgacctgc	atgaaaaaagc	ctgaactcac	cgcgacgtct	gtcgagaagt	6780
ttctgatcga	aaagttcgac	agcgtctccg	acctgatgca	gctctcgag	ggcgaagaat	6840
ctcgtgctt	cagtttcgat	gtaggaggc	gtggatatgt	cctgcgggta	aatagctgcg	6900
ccgatggtt	ctacaaaagat	cgttatgtt	atcggcactt	tgcacatggcc	gctcgtcccg	6960
ttccggaaat	gcttgcacatt	gggaaattca	gcgagagcct	gacctattgc	atctcccgcc	7020
gtgcacaggg	tgtcacgttg	caagacctgc	ctgaaaccga	actgcccgt	gttctgcagc	7080
cggtcgcgg	ggccatggat	gcatcgctg	cggccgatct	tagccagacg	agcgggttcg	7140
gccattcgg	accgcaagga	atcggtaat	acactacatg	gctgtatttc	atatgcgcga	7200
ttgctgatcc	ccatgtgtat	caactggcaaa	ctgtgatgga	cgacaccgtc	agtgcgtccg	7260
tcgcgcaggc	tctcgatgag	ctgatgctt	ggccgagga	ctgccccgaa	gtccggcacc	7320
tcgtgcacgc	ggatttcggc	tccaacaatg	tcttgacgg	caatggccgc	ataacagogg	7380
tcatttgcgt	gagcgaggcg	atgttcgggg	attcccaata	cgaggtcgcc	aacatcttct	7440
tctggaggcc	gtgggtggct	tgtatggac	agcagacgcg	ctacttcgag	cgaggagcattc	7500
cggagcttgc	aggatcgccg	cggctccggg	cgtatatgt	ccgcatttggt	cttgaccaac	7560
tctatcagag	cttgggttgc	ggcaatttcg	atgatgcagc	ttggggcag	gttgcgtgcg	7620
acgcaatcgt	ccgatccgg	gcccggact	tccggcgtac	acaatgcgc	cgcagaagcg	7680
cggccgtctg	gaccgtatgc	tgtgttagaa	tactcgccg	tagtgaaac	cgacgcccc	7740
gcaactcgtcc	gagggcaaaag	gaatagagta	tgcaaaaatt	gatgatctat	taaacaataa	7800
agatgtccac	taaaatggaa	gttttctg	tcataacttt	ttaagaagg	tgagaacaga	7860
gtacctacat	tttgaatgg	aggattggag	ctacgggggt	gggggtgggg	tgggattaga	7920
taaatgcctg	ctcttactg	aaggctctt	actattgtt	tatgataatg	tttcatagtt	7980
ggatatacata	atttaaaca	gcaaaaccaa	attaaggcc	agctcattcc	tcccactcat	8040
gatctataga	tctatagatc	tctcgggg	tcattgttt	tctcttgatt	cccaactttgt	8100
ggttcttaat	actgtgggtt	ccaaatgtgt	cagtttcata	gcctgaagaa	cgagatcagc	8160
agcctctgtt	ccacatacac	ttcatttca	gtattgttt	gccaagttct	aattccatca	8220
gattttatgt	gacatcttta	ttgtttaata	gatcatcaat	ttctgcatac	taagcttac	8280
gataccgtcg	acctcgagg	ggggcccggt	accagctt	tgtccctt	agtgagggtt	8340
aattgcgcgc	ttggcgtaat	catggtcata	gtgtttcc	gtgtgaaatt	gttatccgct	8400
cacaattcca	cacaacatac	gagccgaaag	cataaaatgt	aaagccctgg	gtgcctaatt	8460
agttagctaa	ctcacattaa	ttgcgttgcg	ctcaactgccc	gctttccagt	cgggaaaccc	8520
gtcgtgccag	ctgcattaaat	gaatcgccca	acgcgcgggg	agaggcggtt	tgcttattgg	8580
gogctttcc	gttcctcgc	tcactgactc	gctgcgtcg	gtcggtggc	tgccggcgagc	8640
ggtatcagct	cactcaaagg	cgttaatacg	gttatccaca	gaatcagggg	ataacgcagg	8700
aaagaacatg	tgagccaaa	gccagcaaaa	ggccaggaaac	cgtaaaaagg	ccgcgttgc	8760
ggcggtttcc	cataggctcc	gccccctga	cgagcatcac	aaaaatcgac	gtcgtactca	8820
gaggtggcga	aacccgacag	gactataaag	ataccaggcg	tttccccctg	gaagctccct	8880
cgtgcgtct	cctgttccga	ccctgcgcgt	taccggatac	ctgtccgcct	tttcccttc	8940
gggaaggctg	gcttttctc	atagctacg	ctgttaggtat	ctcagttcg	tgttaggtcg	9000
tcgctcaag	ctgggtgtg	tgcacgaacc	ccccgttcag	cccgaccgc	gcccattatc	9060
cggtaactat	cgtcttgcgt	ccaaccgg	aagacacgc	ttatgcac	tggcagcagc	9120
cactggtaac	aggattagca	gagcgaggta	tgttaggcgg	gctacagagt	tcttgaagtg	9180
gtggccta	acggctaca	ctagaaggac	atgatgggt	atctgcgtc	tgctgaagcc	9240
agttaccc	ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	ccgctggtag	9300

cgggtggttt	tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaggat	ctcaagaaga	9360
tcctttgatc	ttttctacgg	ggtctgacgc	tcagtggAAC	gaaaactcac	gttaaggat	9420
tttggtcatg	agattatcaa	aaaggatctt	cacctagatc	ctttaaatt	aaaaatgaag	9480
ttttaaatca	atctaaagta	tatatgagta	aacttggct	gacagttacc	aatgcttaat	9540
cagtggggca	cctatctcg	cgatctgtct	atttcgttca	tccatagtg	cctgactccc	9600
cgtcgtgttag	ataactacga	tacgggaggg	cttaccatc	ggccccagt	ctgcaatgat	9660
acccgcgagac	ccacgctcac	cggtccaga	tttatcagca	ataaaccagc	cagccggaag	9720
ggcccgagcgc	agaagtggtc	ctgcaactt	atccgcctcc	atccagtcta	tttaatttttgc	9780
ccgggaact	agagtaagta	tttgcgcagt	taatagtttgc	cgcaacgttgc	ttgcatttgc	9840
tacaggcatac	gtgggtgtcac	gtctgtcg	tgttatggct	tcattcagct	ccggttccca	9900
acgatcaagg	cgatcacat	gatccccat	gttgcgaaa	aaagcggtt	gctccttcgg	9960
tcctccgatc	gttgcagaa	gtaagtggc	cgcagtgta	tcaactcatgg	ttatggcagc	10020
actgcataat	tctcttactg	tcatgcacatc	cgttaagatgc	ttttctgtga	ctgggtgagta	10080
ctcaaccaag	tcattctgag	aatagtgtat	gccccgaccg	agttgtctt	gcccgccgtc	10140
aatacggat	aataccgcgc	cacatagcag	aactttaaaa	gtgctcatca	ttggaaaaacg	10200
ttcttcgggg	cgaaaactct	caaggatctt	accgctgttgc	agatccagtt	cgatgttaacc	10260
cactcgtgca	cccaactgat	tttcagcatac	ttttactttc	accagcgttt	ctgggtgagc	10320
aaaaacacgga	aggcaaaatg	ccgcaaaaaaa	gggaataagg	gccccggccgtc	10380	
actcataactc	ttcctttttc	aatattatttgc	aatattatttgc	cagggttatt	gtctcatgag	10440
cggatacata	tttgaatgt	tttagaaaaaa	taaacaaata	gggggtccgc	gcacatttcc	10500
ccgaaaagtgc	ccac					10514

<210>	53					
<211>	168					
<212>	DNA					
<213>	Homo sapiens					
<400>	53					
cgaggacacag	tactccgctc	ggaggacagt	actccgctcg	gaggacagta	ctccgctcg	60
ggacacat	ccgctcgagg	gacagtactc	cgatccgctcg	actctagtcc	gctcgagg	120
cagactccg	ctcggaggac	agtactccgc	tcggaggaca	gtactccg		168

<210>	54					
<211>	22					
<212>	DNA					
<213>	Artificial Sequence					
<220>						
<223>	Description of Artificial Sequence: Synthetic primer					
<400>	54					
gagaacccac	tgcttactgg	ct				22

<210>	55					
<211>	20					
<212>	DNA					
<213>	Artificial Sequence					
<220>						
<223>	Description of Artificial Sequence: Synthetic primer					
<400>	55					
caagatcgcc	ctcgcccttg					20

```

<210> 56
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 56
atatggattcc gcagtgaaaa tgtctcataaa tgg 33

<210> 57
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 57
aattggatcc tcaatcagtc catcacctgg 30

<210> 58
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 58
cgtcttagaga gtgtgtcctg tcagaagaac 30

<210> 59
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 59
atgcggccgc actcctgacc tcaggtgatc c 31

<210> 60
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 60
cgtcttagaga agacagtcaa ggtcgtgac 29

```

```

<210> 61
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 61
gcgccggcgc gtctaaccta ggagccac 29

<210> 62
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 62
gctctagagg acgaatgcc a agatctga 28

<210> 63
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 63
caagcggccg ccagtgtat ggatatctg 29

<210> 64
<211> 5218
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant
      DNA sequence

<400> 64
gccggagcgg tcgagttctg gaccgacccgg ctcgggttct cccgggactt cgtggaggac 60
gacttcggc gtgtggtcgg ggacgcacgtg accctgttca tcagegcggc ccaggaccag 120
gtggtgccgg acaacaccct ggcctgggtg tgggtgcgcg gcctggacga gctgtacgcc 180
gagttggcgg aggtcgtgtc cacgaacttc cgggacgcct ccggggccgg catgaccgag 240
atcggcgagc agccgtgggg gcgggagttc gcctggcgcg accccggccgg caactgcgtg 300
cacttcgtgg ccgaggagca ggactgacac gtgctacgag atttcgattc caccggccgcc 360
ttctatgaaa ggttgggctt cgaaatcggtt ttccgggacg ccggctggat gatcctccag 420
cgcggggatc tcatgctggta gttttcgcc cacccaaact tgtttattgc agcttataat 480
ggttacaaat aaagcaatag catcacaaat ttcacaaata aagcattttt ttcaactgcat 540
tctagttgtg gtttgtccaa actcatcaat gtatcttattc atgtctgtat accgtcgacc 600
tctagctaga gcttggcgta atcatggtca tagctgttcc ctgtgtgaaa ttgttatccg 660

```

tcacacaatcc cacacaacat acgagccgga agcataaagt gtaaaaggcctg ggggtgcctaa
ttagtgagct aactcacatt aattgcgttg cgctcaactgc ccgctttcca gtcgggaaac 780
ctgtcgcc agtcatttta atgaatccgc caacgcgcgg ggagaggcgg ttgcgtatt 840
ggcgctttt ccgcttcctc gctcaactgc tcgctgcgtt ccgtcggtcg gctgcggcga 900
ggcggtatcag ctcaactaaa ggcggtaata cggttatcca cagaatcagg ggataacgcga 960
ggaaagaaca tggagcaaa aggccagca aaggccagga accgtaaaaaa ggccgcgttg 1020
ctggcgcccc tccataggct ccgccccct gacgagcatc aaaaaaatcg acgctcaagt 1080
cagaggtggc gaaaccgcac aggactataa agataccagg cgccccccccc tgaaagctcc 1140
ctcggtcgct ctccgttcc gaccctgcgg ctaccggat acctgtccgc ctttctccct 1200
tcgggaagcg tggcgcttcc tcaatgetca cgctgttaggt atctcagttc ggttaggtc 1260
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccggccctt ctgcgcctt 1320
tcggtaact atcgtcttga gtccaaaccgg gtaagacacg acttatcgcc actggcagca 1380
gccactggta acaggattag cagagcgggg tagttaggcg gtgctacaga gttcttgaag 1440
tggtgtccca actacggcta cactagaagg acagtattt gatatctgc tctgctgaag 1500
ccagttaccc tggaaaaaaag agttggtagc tcttgcattcc gcaaacaac caccgctgg 1560
agcgggtggtt ttttggtttca caagcggcgg attacgcgcg gaaaaaaaaagg atctcaagaa 1620
gatccttga tctttctac ggggtctgac gtcagtgga acgaaaaactc acgtaaggg 1680
attttggtca ttagatttca aaaaaggatc ttacccataga tcctttttaaa taaaaatgaa 1740
agttttaaat caatctaaag tatataatgggaa taaaacttggg ctgacagtta ccaatgttt 1800
atcagtgggg caccatctc acgatctgt ctatctcgat catccatagt tgccctgactc 1860
cccgctgtgt agataactac gatacgggg ggttccat acgtggccatc tgctgcatt 1920
ataccggag acccacgctc accggctca gatttacatca caataaaacca gccagccgg 1980
agggccgagc gcagaagtgg tcctgcacta ttatccgcct ccatccagtc tattaatttg 2040
tgccgggaag cttagtgcgg tagttcgcc gttaatagtt tgccaaacgt tggtgcatt 2100
gctacaggca tgggtgttca acgctcgatc ttgggtatgg cttcattcag ctccgggttcc 2160
caacgatcaa ggcgagttac atgatcccccc atgatgtca aaaaagcggt tagctcttc 2220
ggtcctccga tgggtgttca aagtaatgg gccgcgtgt tatcactcat gtttatggca 2280
gcactgcata atttcttac tggatgcggca tccgtaaat gctttctgt gactggtgag 2340
tactcaacca agtcattctg agaataatgtt atgcggcgac cgagtgtctc ttgcccggcg 2400
tcaataacggg ataataccgc gcccataatgc agaactttaa aagtgtctcat cattggaaaa 2460
cgttcttcgg gggaaaact ctcaaggatc ttaccgcgt tgagatccag ttgcgtgtaa 2520
cccactcgat caccctactg atcttcagca tctttactt tcaccagcg ttctgggtga 2580
gcaaaaacag gaaggcaaaa tgccgaaaaa aagggaataa gggcgacacg gaaatgttga 2640
atactcatac tcttcctttt tcaatattat tgaagcattt atcagggttta ttgtctcatg 2700
agcggataca tatttgaatg tatttagaaa aataaaacaaa taggggttcc ggcgcacattt 2760
ccccgaaaaag tgccacactg cgtcgacggc tcgggagatc cgatgttttca tgagttact 2820
ttgtatcccc accccccctc gaggatcgat ccgctcgagg gacagatctc cgctcgagg 2880
acagtactcc gctcgagga cagtaatccg ctcgaggaca gtactccgcg cggaggacac 2940
tactccgatc cgtcactct agtccgcgt gaggacagta ctccgcgtgg aggacagtac 3000
tccgctcgat ggacagttact ccgctcgagg acagatctcg attcgaggac attttgacac 3060
ccccataata ttttccaga attaacatgt taaattgcatt ctctgttca agagttccct 3120
atcactcttca ttaatcacta ctcacagtaa ctcacactcc tgcccaagct tggttaccacc 3180
atggacccttcc aacgcgtgg gaaagttaaa gatgtgttca agtacttggg tgacgcgttg 3240
ggttacatcg aacttggatct caacagcggt aagatccctt agagttttcg ccccgaaagaa 3300
cgttttccaa tggatgttca ttttaaattt ctgtatgtt ggcgggttattt atcccgatgtt 3360
gacggccggc aagagcaact cgggtccgcg atacactatt ctcagaaatga ttgggtttag 3420
tactcaccag tcacagaaaaa gcatcttacg gatggcatga cagtaagaga attatgcagt 3480
gctggccatccaa ccatgagtga taacactgcg gccaacttac ttctgacaaac gatcgaggaa 3540
ccgaaggagc taaccgcgtt tttgcacaaat atgggggatc atgtaactcg ctttgcgtt 3600
tgggaaccgg agctgtatgt gacccataccat aacgacgacgac gtgacaccac gatgcctgt 3660
gcaatggca aacgttgcg caaacttata actggcgaaac tacttactct agcttcccg 3720
caacaattaa tagactggat ggaggccgtt aaagttgcgtt gaccacttct ggcgcgtgg 3780
cttccggctg gctggtttat tggatgttca tctggagccg gtgaggttgg gtctcgccgtt 3840
atcattgcag cactggggcc agatggtaag ccctcccgta tcgttagttt ctacacgcac 3900
gggagtcagg caactatggta tgaacgaaat agacagatcg ctgagatagg tgccctactg 3960
attaagcatt ggtatctag agagggcccg tttaaaccgg ctgatcagcc tcgactgtgc 4020
cttctagttt ccagccatct gttgttgcctt ccctcccgat gcttccctt accctggaaag 4080
gtgccactcc cactgtccctt tcctaataaa atgagggaaat tgcatcgat tgctcgatgtt 4140

ggtgtcattc tattctgggg ggtggggtgg ggcaggacag caagggggag gattgggaag	4200
acaatagcag gcatgctggg gatgcgggtgg gctctatggc ttctgaggcg gaaagaacca	4260
gctgggctc tagggggtat ccccacgcgc cctgtacgcg cgcattaagc gcggcgggtg	4320
tggtggtac ggcagcgtg accgctacac ttgccagcgc cctagccccc gtcctttcg	4380
ctttcttccc ttcccttctc gccacgttcg ccggcttcg ccgtcaagct ctaaatcg	4440
gcatccctt agggttccga tttagtgcgt tacggcacct cgaccggaaa aaacttgatt	4500
agggtatgg ttacgttagt gggccatcgc cctgatagac ggttttcgc ctttgacgt	4560
tggagtccac gtctttaat agtggactct ttgttccaaac tggaacaaca ctcaacccta	4620
tctcggtcta ttctttgtat ttataaggga ttggggat ttccggctat tggttaaaaa	4680
atagctgat ttaacaaaaa ttaacgcga attaattctg tggatgtgt gtcagttagg	4740
gtgtggaaag tccccaggct ccccaggcg gcagaagtat gcaaagcatg catctcaatt	4800
agtcaaac caggtgtgaa aagtccccag gtcggccagc aggccagaat atgcaaaagca	4860
tgcatacaa ttatcagca accatagtcc cggccctaaac tccggccatc cggccctaa	4920
ctccgcccag ttccgccccat tctccgcccc atggctgact aattttttt atttatgcag	4980
aggccgaggc cgccctcgcc tctgagctat tccagaagta gtgaggaggc tttttggag	5040
gcctaggctt ttgcaaaaag ctccggggag ctgtatatac catttcgga tctgtatcagc	5100
acgtgtgac aattaatcat cggcatagta tatcggcata gtataatacg acaaggtag	5160
gaactaaacc atggccaagt tgaccagtgc cgttccgggtg ctcaccgcgc ggcacgtc	5218

<210> 65
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 65
gggttaccaa agaaggttagg ctgagggg

28

<210> 66
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 66
cgggatccgt acaagacagt tagctagttg gc

32

<210> 67
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 67
gggttacccat tcagcaggcc actacaggac tctc

34

```

<210> 68
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

<400> 68
gatcgccggcc gctgtgctcg acattggtgg cc          32

<210> 69
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

<400> 69
ggggtaacctt tgggtgtc tt aaaaaatgag c          31

<210> 70
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

<400> 70
gatcgccggcc gcctcatgaa caatgaaatc tcc          33

<210> 71
<211> 2379
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      Nurrl targeting sequence

<400> 71
cgtacaagac agtagctag ttggcaaaac caaggagaat tctgtgacaa gggaaactca 60
ggacaaatcc tgctaggcag ttctggaggg gaagcgccct gcgcctgcag tactgacctg 120
tgaccatagc cagggcagca atgcaggaga agcagaaaat gtcgatgttc atattctgca 180
agttggagga gaattcaaca atggaatcaa tccattcccc aaagccacga acgcattgca 240
acctgtgcaa gaccacccc ttgcaaaaga tgagtttacc ctccactggg ttggacctgc 300
aattaatacc aaagagagag aggggagaaa aagagagaga gaaaagctaa acaactaatt 360
acttgaagag caataaatga gggatttaag agtcacctaa ttactgaaga gttaattaaa 420
atgttagatcc agtggacctt gaaagggttt aatttcataaa caatcaagaa cgccccctat 480
cccacctcca ttccattcat ccaaggcttc tgggctcgct ttttttgtt 540

```

ttctttcttt ttcttccttt ctccgacttc catttcctat tctgtctttt tctctacccc 600
 accctctggc ttcccttcct ccctttcttt tccttccttg atttctctca cagcctccct 660
 ggattgtctc ctccctccct tattacctgt atgctaatacg aaggacaaac agttctaaga 720
 aagctgattc aaaaagcagg tcttggtcgg ctttgggcag gtctgcaag ccagggatct 780
 tctctgccc gccccggatg atctccatgg agccagtcag gagatcatag aattgctgga 840
 tatgctgggt gtcatctcca ctcattgtat agtcagggtt cgccctgaaac tggaaatttca 900
 ttttaaaag cactaatga gtttctctaa aatatataaac ccgtgaaatt gctaaccccc 960
 ttcttaatag gggagccagg ttttataaac aattaacccct ctctctgacc agtaaggaaa 1020
 tttagatgttcc cggggcaatt aattcaattt gggatgggtc tacgaggtcg ctgctttaaa 1080
 tatgttaattt gtcatctcca tacagactaa atacagtgc atccagccct gggaaacggtt 1140
 cactcttac tccacaaaaac aattgacatg ttagtattca tgcttagtccc agagtgggcc 1200
 tcgagctgag acggccctttt gcaaatactta aaaaaattgc agtccctacc aggtttgctc 1260
 aacataacttc ctccaaattaa agccgcata ttttcacag ctggagtaaa aggtatcgac 1320
 cacttgatcc cccacccccc tggcatctt acaccctctc ctcccttcc ctgacccccc 1380
 gacagcccccc acatccttcc ccagcacagt gccaggaagg atgaggatta aagcatctat 1440
 ctctatggag gactggaaat aaagttagatt gggaaaggggg agagttgca cagccctgg 1500
 atttcatggt agaggcactg gtgtatatgg aagaggagggg gcgatgggggt agtggggtag 1560
 tgggaattca cacagcttca aaccctgtgg cgtgtccctc tcctccttcc tccatctgt 1620
 cctcatgtct atgccccccag ctacatggat ttgtctgaag atctgaatct tgaggccctg 1680
 gccagagctg cgaggcatat acagccttgc ttgccttct ttacccccc tgaatctgag 1740
 agttaatgac ggtatgtgggg aggggtctg cccatctgtc gggttgcaca catgatatcc 1800
 ccccccggcag ctcttaccc tggaaatagtc caggctggtc atagccgggt tggagtcgac 1860
 atggggccctg acgagggcac tgatcagact caccgggggc gaagggggag agggctcctg 1920
 tgggcttcc ggttgcagg gcaaacgacc tctccggcct tttaaactgt ctgtgcgaac 1980
 cactgcaaag gaagagccct gttagcggcgg ctttccgag cccaggccca gctgctgcct 2040
 cggcccttcc cgggggaagg cccgcggcgc ggggcaccag gctgagccgc tgagggccccc 2100
 agtgcgtgtt aagccttcac tgactagaag cattaaaaaa tgcgggggta ttttatgtct 2160
 tcctccaaat gggtcgata gttaaaggag agaagggcct ggcggcttcc tctagggaaag 2220
 gccggggcaag caggcagctg cagggtctg gaggccatac tgagggggag tccggagatcc 2280
 ccagcaccgg gaagtggaaac gtgatgctgg agtatgagca gttgtttccaaaggcccaa 2340
 actggagggtt cggcagctcc cctcagccca ctttctttt 2379

<210> 72
 <211> 8426
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 pCDGal4-DBD-Nurrl sequence

<400> 72
 tctctggctta actagagaac ccactgctta ctggcttatac gaaattaata cgactcacta 60
 tagggagacc caagctggct agcgtttaaa ctaagcttg ccaccatgaa gctactgtct 120
 tctatcgaaac aagcatgca tatttgcga cttaaaaaagc tcaagtgtct caaagaaaaaa 180
 ccgaagtgcg ccaagtgtct gaagaacaac tggaggtgtc gctactctcc caaaacccaa 240
 aggtctccgc tgactaggc acatcgaca gaagtggaaat caaggctaga aagactggaa 300
 cagctatttc tactgatttt tcctcgagaa gaccttgaca tgatggatct 360
 ttacagagata taaaagcatt gttAACAGGA ttattttgtac aagataatgt gaataaaagat 420
 gccgtcacag atagattggc ttcaagtggag actgatatgc ctctaacatt gagacagcat 480
 agaataatgtt cgcacatcatc atcgaaagag agtagtaaca aaggtaaaag acagttgact 540
 gtatcgccgg gtaccaaaaga aggttaggctg aggggagctg ccgaccctcc agtttgcgcc 600
 ttttagggaaac cactgctcat actccagcat cacgttccac ttccgggtgc tggggatctc 660
 cgactcccccc tcagttatggc ctccagggacc ctgcagctgc ctgcttgcctt ggccttccct 720
 agagaaagcc gccaggccct tctctccctt aactatacga cccatttggaa ggaagacata 780
 aaataacccc gcattttta atgcttcttag tcagtgaaagg cttaacaagc actggggccc 840
 tcagccgctc agcctgggtgc cccggccctg cggccttccc cggggaggga cccggaggc 900

agctgggcct gggctcgaa aagcggcgct aacagggctc ttcctttgca gtggttcgca 960
 cagacagttt aaaaggccgg agaggtcggt tccctcgaa accgaagagc ccacaggagc 1020
 cctctcccccc ttgcggcccg gtgagtcgtc tcagtgcct cgtagggcc catgtcgact 1080
 ccaaccggc tatgaccagg ctggactatt ccagggtaaag aagctggcg gggggatattc 1140
 atgtggacaa accgacagat gggcaggacc cctcccccaca tccgtcatta actctcagat 1200
 tcaacggggg taaagaaaagg caagcaaggc tgtatatgcc tcgcagctct ggccaggggc 1260
 tcaagattca gatcttcaga caaatccatg tagctggggg catagacatg aggacaggat 1320
 ggaggaagga ggagaggggac acgcccacagg gttgaagct gtgtgaattc ccactacccc 1380
 actacccat cgcggccctt cttccatata caccatgc tctaccatgaa aatccagggg 1440
 ctgtgcacaaac tctcccccctt cccaaatctac ttatttccca gtcctccata gagatagatg 1500
 cttaatcct catccttcctt ggcactgtgc tggggaagga tgtggggct gtctgggggt 1560
 cagggaaaggg aaggagaggg tgtaaagatg ccagtggggt gggggatcaa gtggtcagat 1620
 cctttactc cagctgtgaa aaatatgcgg gctttaattt gaggaaagtat gttgagcaaa 1680
 cctggtaggg actgcaattt tattaagatt tcaaaaaggg cgtctcagct cgaggcccac 1740
 tctggacta gcatgaatac taacatgtca attgtttgt ggagataaga gtgaacgttt 1800
 cccagggctg gatggcactg tatttagtct gtatggaaat gacaatttac atatttaaag 1860
 cagcgacccctc gtacgaccat ccctaatttga attaatttgc ccggaaacatc taatttcctt 1920
 actggtcaga gagagggtta attgttataa aaacctggct cccctattag aaacgggggt 1980
 agcaatttca cgggttatattt attttagaga acctcattaa gtgctttta aaatgaaatt 2040
 ccagttccag gegaaccctg actatcaaattt gagtggagat gacaccacg atatccagca 2100
 attctatgtat ctcctgactg gtcctcatgg gatcatccgg ggctggcgag agaagatccc 2160
 tggcttcgca gacctgccc aagccgacca agacctgtttt tttgaatcgt ctttctttaga 2220
 actgtttgtc cttcgatttag catacaggtt ataaggggagg gaggagacaa tccaggagg 2280
 ctgtgagaga aatcaagaaaa gggggggggaa gaaaccacg ggtggggtag 2340
 agaaaaaagac agaataaggaa atggaagtcg gagaaggaa gaaaaaagaaa gaaaacaaaa 2400
 aaagacgaga agaagcgagc ccagaagccct tgatgaatg gaatggaggt gggatagggg 2460
 gcggttttgc ttgttatgaa attaaaccctt ttcaaggtcc actggatcta catttttaatt 2520
 aactcttcag taatgttgactt ctatctttaat ccctcatttt ttgccttca agtaatttagt 2580
 tggtttagt tttctctctt ctttttcttcc ctttctcttc tttgttataatttgcaggtc 2640
 caaccctgtt ggggttaaac tcatcttttgc caatgggggtt gtccttcaca gggttcaatg 2700
 cgttcgtggc tttggggat ggattgttttgc cattgttggaa ttcttccttca acttgcagaa 2760
 tatgaacatc gacatttctg ctttcttcgtt cattgttgc cttgttatgg tcacaggtca 2820
 gtactgcagg cgccggcgc ttcccttcca gaaactgccta gcaggatttgc ttctgagttt 2880
 cccttgcac agaattctcc ttgggtttgc caactagcta actgtcttgc acggatccac 2940
 tagtaacggc cgccagtgttgc ctggatttgc gcaagatattcc atcacactgg cggccgcctg 3000
 agtcttaggg gcccgtttaa acccgctgtat cggcctcgac tgccttctt agttgcacgc 3060
 catcttttgtt ttggcccttcc cccgttgcctt ctttgcaccc ggaagggttgc actccactg 3120
 tcctttcttataaaaatgag gaaatttgcatttgc cgcattgtctt gaggtaggtt catttttattt 3180
 tgggggggtgg ggtggggcag gacagcaagg gggaggattt ggaagacaaat agcaggcatg 3240
 ctggggatgc ggtgggttctt atggcttctg aggcggaaag aaccagctgg ggctcttaggg 3300
 ggtatccccca cgcggccctgtt agcggccgtat taagcgcggc ggggtgtggg gttacgcgc 3360
 gcggtaccgc tacacttgcg acgcggccctag cggccgcctt ttcgttccctt ttcccttcc 3420
 ttctcgccac gttcgccggc tttcccttgc aagctctaaa tcgggggttgc cttttaggg 3480
 tccgatttag tgctttacgg caccttgcacc cccaaaaactt tgatttaggtt gatggttcac 3540
 gtagtggggcc atcgccctgtt tagacgtttt ttcccttgc gacgttggag tccacgtttt 3600
 ttaatagtgg actcttgcgtt caaaacttgcgaa caacacttca cccatcttcg gtttatttctt 3660
 ttgatttata agggatttttgc cggatttgcg ctttgcgtt aaaaaatgag ctgattttaac 3720
 aaaaatttaa cgcgaattaa ttctgtggaa tttgtgttcgat tttaggtgtt gaaagtcccc 3780
 aggctccccca gcaggccggaa gttatgcgttcc aatttagtgcg caaccagggtt 3840
 tggaaaggccc ccaggctccc cggccggcag aagtatgcgaa agcatgcatttcaatttagt 3900
 agcaaccata gttccggccccc taacttgcgc catccggccc ctaacttgcgc ccagttccgc 3960
 ccatttgcgtt cccatggctt gactaattttt ttttatttgc gcaaggccgc aggccgcctt 4020
 tgccttcgttgc ctatccaga agtagtgcgtt aggtttttttt ggaggccctag gtttttgc 4080
 aaagctccccca ggagcttgcgtt tatccattttt cggatctgtat cggcactgtca tgaaaaagcc 4140
 tgaacttccacc gcgacgtctg tcgagaagtt tctgtatgcgaa aagttcgaca gctgtccgc 4200
 cctgtatgcgttgc ctctcgagg gcaagaatc tcgtgttttgc agttcgatg taggaggggc 4260
 tggatatgtc ctgcgggtaa atagctgcgc cgtatggtttca tacaagatc gttatgtttt 4320
 tcggcactttt gcatcgccgc cgctcccgat tccggaaagtttgc cttgacatttgc gggaaatttgc 4380

cgagagcctg acctattgca tctcccgccg tgcacagggt gtcacgttgc aagacctgcc 4440
 tggaaaccgaa ctgcccgtg ttctgcagcc ggtcgcggag gccatggatg cgatcgctgc 4500
 ggccgatctt agccagacga gcggggtcgg cccattcggaa ccgcaaggaa tcggtaata 4560
 cactacatgg cggtgatttca tatgcgcgat tgctgatccc catgtgtatc actggcaaac 4620
 tgtgtatggac gacaccgtca gtgcgtccgt cgogcaggct ctcgatgagc tgatgctttg 4680
 ggccgaggac tgccccgaag tccggcacct cgtgcacgcg gatttcggct ccaacaatgt 4740
 cctgacggac aatggccgca taacagcggt cattgactgg agcggcgcg tttcgaaaa 4800
 ttcccaatac gagggtcgcca acatcttctt ctggaggccg tgggtggctt gtatggagca 4860
 gcagacgcgc tactcgagc ggaggcatcc ggaggttgcg ggategcgcg ggctccgggc 4920
 gtatatgtc cgattggtc ttgaccaact ctatcagagc ttgggttgcg gcaatttcga 4980
 tgatgcagct tgggcgcagg gtcgatgcgca cgcacatcgct cgatccggag ccggactgt 5040
 cgggcgtaca caaatcgccc gcagaagcgc gcccgtctgg accgatggct gtgtagaagt 5100
 actcgccat agtgaaacc gacgccccag cactcgccg agggcaaagg aatagcacgt 5160
 gctacgagat ttgcatttca ccgcgcctt ctatgaaaagg ttgggttgcg gaatcgcccc 5220
 ccgggacgccc ggctggatga tcctccagcg cggggatctc atgctggagt tttcgcccc 5280
 ccccaacttgc ttatttgcag cttataatgg ttacaaataaa agcaatacgca tcacaaattt 5340
 cacaataaaa gcattttttt cactgcattt tagttgtggg ttgtccaaac tcatcaatgt 5400
 atcttatcat gtctgtatac cgatcgaccc tagcttaggc ttggcgtaat catggtcata 5460
 gctgtttcct gtgtaaaatt gttatccgct cacaattcca cacaacatac gagccgaaag 5520
 cataaaagtgt aaaggctggg gtgcctaattt agtgagctaa ctcacattaa ttgcgttgcg 5580
 ctcactgccc gctttccagt cgggaaacct gtcgtgcacag ctgcattaaat gaatcgccca 5640
 acgcgcgggg agaggcggtt tgcgtattgg gcgccttcc gcttcctcgc tcactgactc 5700
 gctgcgtcgt gtcgttcggc tgcggcgcggc ggtatcagct cactcaaagg cggtatcacg 5760
 gttatccaca gaatcagggg ataacgcagg aaagaacatgt tgagcaaaag gccagcaaaa 5820
 ggccaggaac cgtaaaaagg cgcgttgct ggcgttttc cataggctcc gccccctga 5880
 cgagcatcac aaaaatcgac gtcgaatgtca gaggtggcga aacccgacag gactataaaag 5940
 ataccaggcg tttccccctg gaagctccct cgtgcgcctt cctgtccga ccctggcgct 6000
 taccggatac ctgtccgcct ttctccctt gggaaagcgtg ggccttccatc atagctcacg 6060
 ctgttaggtat ctgcgttccgg tgcgttgcgt tgcgttccaaag ctgggtgtg tgacgaaacc 6120
 ccccggtcag cccgaccgcg ggccttccatc cgtaactat cgtcttgagt ccaaccgggt 6180
 aagacacgcac ttatcgccac tggcagcgcg cactggtaac aggattagca gagcggggta 6240
 tgcgttccgg tgcgttccatc ttgcgttccatc tgcgttccatc tgcgttccatc 6300
 agtatttggt atctgcgtc tgctgaagcc agttacccctt ggaaaaaagag ttggtagctc 6360
 ttgcgttccggc aaacaaaacca ccgcgttccatc cgtttttttt gtttgcagc agcagattac 6420
 ggcgttccggc aaaggatctc aagaagatcc ttgcgttccatc tgcgttccatc tgcgttccatc 6480
 gtggaaacgaa aactcacgtt aagggtttt ggtcatgaga ttatcaaaaa ggatcttcac 6540
 ctgcgttccatc ttggatccatc aatgcgttccatc taaatcaatc taaatcaatc taaatcaatc 6600
 ttgggttccatc agttaccaat gcttaatccatc tgaggcaccc atctcagcga tgcgttccatc 6660
 tgcgttccatc atagggttccatc gactcccccgt cgtgttagata actacgatac gggagggtt 6720
 accatctggc cccagtcgtc caatgatacc gcgagaccca cgctcaccgg ctccagattt 6780
 atcagcaata aaccagccag ccggaaaggcc cgagcgcaga agtggcttcgca cacttttac 6840
 cgcgttccatc cagtcttata attgttccgg ggaagctaga gtaagtagtt cgccagttaa 6900
 tagtttgcgc aacgttggcc ccattgttccatc aggcgttccatc gtttgcgttccatc 6960
 tttttttccatc ttgcgttccatc gtttgcgttccatc atcaaggccg gtttgcgttccatc 7020
 gtcgttccatc gtcgttccatc gtttgcgttccatc tttttttccatc tttttttccatc 7080
 agtggatccatc ctgcgttccatc tggttccatc tttttttccatc tttttttccatc 7140
 aagatgtttt ttcgttccatc tggttccatc tttttttccatc tttttttccatc 7200
 ggcgttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 7260
 tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 7320
 gtcgttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 7380
 tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 7440
 aataaggccg acacggaaat ttgcgttccatc tttttttccatc tttttttccatc 7500
 cttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 7560
 acaaataaggccg ttgcgttccatc tttttttccatc tttttttccatc tttttttccatc 7620
 agatctcccg atcccttccatc ttgcgttccatc tttttttccatc tttttttccatc 7680
 agccgttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 7740
 taagttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 7800
 cttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 7860

agttattaat agtaatcaat tacggggtca ttagttcata gcccataatat ggagttccgc 7920
 gttacataac ttacggtaaa tggccccct ggctgaccgc ccaacgaccc ccccccattg 7980
 acgtcaataa tgacgtatgt tcccatagta acgccaatag ggacttcca ttgacgtcaa 8040
 tgggtggagt attacggta aactgcccac ttggcagttac atcaagtgt a tcatatgcc 8100
 agtaccccc ctattgacgt caatgacggt aaatggcccg cctggcatta tgcccaatg 8160
 atgacccat gggactttcc tacttggcag tacatctac tattagtcat cgctattacc 8220
 atggtgatgc ggtttggca gtacatcaat gggcgtggat agcggttga ctcacgggga 8280
 ttccaatgc tccacccat tgacgtcaat gggagttgt tttggcacca aaatcaacgg 8340
 gactttccaa aatgtcgtaa caactccgcc ccattgacgc aaatggcgg taggcgtgt 8400
 cggtggagg tctatataag cagagg 8426

<210> 73
 <211> 9701
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 pKI-Gal4-DBD-Nurrl sequence

<220>
 <221> modified_base
 <222> (1518)...(1518)
 <223> any nucleotide

<400> 73
 tgcacgttc aaaagcgcac gtctggcg 60
 cgacgtcat gaaaaaggct gaactcaccg cgacgtctgt cgagaagttt ctgatcgaaa 120
 agttcgacag cgtctccgac ctgatgcgc ttcggagg 180
 gttcgatgt aggaggcg 240
 acaaagatcg ttatgtttat cggcacttg catcgccgc gtcggatt ccggaaatgc 300
 ttgacattgg ggaattcage gagagcctga cctattgcat ctccggcgt gcacagggtg 360
 tcacgttgca agacgtgc 420
 ccatggatgc gatcgctgcg gccgatcta gccagacgag cgggttcggc ccattcgac 480
 cgcaaggaaat cggtaatac actacatggc gtgatttcat atgcgcgatt gtcgtatccc 540
 atgtgtatca ctggcaact gtgatggacg acaccgtc 600
 tcgatgagct gatgtttgg gccgaggact gccccgaatg cccgcaccc gtgcacgcgg 660
 atttcgctc caacaatgtc ctgacggaca atggccgat aacagcggtc attgactgga 720
 gcgaggcgat gttcgggat tcccaatacg agtgcggca 780
 ggttggctt 840
 tatggagcag cagacgcgt acttcgagcg gaggcatccg gagcttgcag
 gatgcgcg gctccggcg 900
 tggttgacgg caatttcgat gatgcagtt gggcgcagg 960
 gatccggagc cggactgtc gggcgtacac aaatcgcccg cagaagcg 1020
 ccgatggctg ttagaagta ctcggcata gtggaaaccg acgccccagc actcgccga 1080
 gggcaagga atgcctgaga aaggaagtga gctgtaaagg ctgagctc tctctgacgt 1140
 atgttagctc tggtagctt cgtcaactcac tttcttgac tcagcatggc aatctgatga 1200
 aatcccaatgtc gtaagtctgc ataaatgtat gatctattaa acaataaaga tgtccactaa 1260
 aatggaaat tttactgtc atactttgtt aagaagggtg agaacagagt acctacattt 1320
 tgaatgaaag gattggagct acgggggtgg ggggtgggggtt gggattagat aaatgcctgc 1380
 tcttactga aggtcttta ctattgtttt atgataatgt ttcatagttt gatatactaa 1440
 tttaaacaag caaaaaccaaa ttaaggccca gtcatttc 1500
 tatacatctc tctgtggatc attgttntc tgcactcactg gaagcttac gataccgtcg 1560
 acctcgagg 1620
 gggcccggtt acccagctt tttccctt agtgagggtt aattgcgcgc
 ttggcgtat catggcata gctgttcctt gtgtgaaattt gttatccgctt cacaattcca 1680
 cacaacatac gagccggaaag cataaaagtgt aaagcctggg gtcctaatg agtggactaa 1740
 ctcacattaa ttgcgttg 1800
 ctgcattaaat gaatcgccca acgcgcgggg agaggcggtt tgcgtattgg ggcgtttcc 1860

gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc tgccgcgagc ggtatcagct 1920
 cactcaaagg cgtaataacg gttatccaca gaatcagggg ataacgcagg aaagaacatg 1980
 tgagaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgc ggcgttttc 2040
 cataggctcc gc(ccccc)ta cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 2100
 aaccgcacag gactataaaag ataccaggcg tttccccctg gaagctccct cgtgcgctct 2160
 cctgtccga ccctgcgcgt tacccgatc ctgtccgcct ttctccctc gggaaagcgtg 2220
 gcgccttc atagctcagc ctgtaggtat ctcaatcgag ttaggtcg tgcgtccaag 2280
 ctggcgtgtc tgcaacgacc ccccgatc ccgcgcgtc ggcgttatac cggttaactat 2340
 cgtcttgagt ccaaccccggt aagacacgcg ttaggtcg ttaggtcg tgcgtccaac 2400
 aggattagca gagcggatc ttaggtcg ttaggtcg ttaggtcg tgcgtccaac 2460
 tacggctaca cttagaaggac agtattttgt atctgcgtc tgcgttaaagcc agttaccttc 2520
 ggaaaaagag ttgttagctc ttgatccggc aacaaaacca ccgcgttgc cgggtgggtt 2580
 tttgttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc 2640
 ttttctacgg ggtctgacgc tcagtggaaac gaaaactcactt gttaaaggat ttgtgtcatg 2700
 agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag tttaaatca 2760
 atctaaagta tatatgagta aacttggtct gacagttacc aatgttaat cagtggcga 2820
 cctatctcag cgatctgtct attcgttca tccatagttt cctgactccc cgtcgtgtag 2880
 ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac 2940
 ccacgcacac cgctccaga ttatcagca ataaaccaggc cagccgaaag ggccgagcgc 3000
 agaagtggtc ctgcaacttt atccgcctcc atccagtcata ttaatgttg ccggaaagct 3060
 agagtaagta gtgcgcgtt taatagttt cgcacgttgc ttgcattgc tacaggcattc 3120
 gtgggtcac gtcgtcggtt tggatggct taattcagct ccgggtccca acgatcaagg 3180
 cgagttacat gatccccat gttgtgcaaa aaagcggta gtccttcgg tcctccgatc 3240
 gttgtcagaa gtaagttggc cgcagttta tcaactcatgg ttatggcagc actgcataat 3300
 tctttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag 3360
 tcattctgag aatagtgtat gcccgcaccc agttgtctt gcccgcgtc aatacggat 3420
 aataccgcgc cacatagcag aactttaaa gtcgtcatca ttggaaaacg ttcttcgggg 3480
 cggaaaactct caaggatctt accgcgttgc agatcggatc cgatgtacc cactcgtgca 3540
 cccaaactgt ctccagcatc ttttacttcc accagcgtt ccgggtgagc aaaaacaggc 3600
 aggcaaaatg ccgcaaaaaaaaa gggaaaatcggc ggcacacggc aatgttgaat actcataactc 3660
 ttcccttttc aatattattt aagcattttt cagggttatt gtctcatgag cggatacata 3720
 tttgaatgtt ttttagaaaaaa taaaacaaaata ggggttccgc gcacatttc ccggaaatgt 3780
 ccacctaaat tgaagcgtt aatattttgt taaaatttcgc gttaaaatttt tgtaaatca 3840
 gtcattttt taaccaatag gccgaatcgc gcaaaatccc ttataatca aaagaataga 3900
 ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 3960
 actccaaacgt caaaggcgcgaaaaccgtct atcaggcgcg tggccacta cgtgaaccat 4020
 caccctaatac aagtttttgc gggtcgaggt ggcgttaaaggc actaaatcgg aaccctaaag 4080
 ggagcccccg atttagagct tgacggggaa agccggcga cgtggcgaga aaggaaaggga 4140
 agaaaagcga aggagcgggc gctagggcgc tggcaagtgt agcgggtcacc ctgcgcgtaa 4200
 ccaccacacc cggccgcgtt aatgcgcgc tacaggcgc gtccttcgc ccattcaggc 4260
 tgcgcaactg ttgggaaggc cgatcggtgc ggccttcgc gtcattacgc cagctggcga 4320
 aaggggatg tgctgcaagg cgattaagtt ggttaacgcg agggtttcc cagtcacgc 4380
 gttgtaaaac gacggccagt gacgcgcgt aatacgcactc actataggc gaattggagc 4440
 tccaccgcgg tgccgcggg ccatgcaggc caccgacatga taagatacat tgatgagtt 4500
 ggacaaacca caactagaat gcaatggaaa aatgtctta tttgtgaaat ttgtgtatgc 4560
 attgttttat ttgttaaccat tataagctgc aataaacaag ttcttgatc gtcgtccat 4620
 gccgagatgt atccccggc cggtcacgaa ctccagcagg accatgtgtat cgcgttctc 4680
 gttgggtct ttgtcgatc cggactggta gtcaggtgt tggttgcgg gcagcgcac 4740
 gggccgcgtc ccatgggggg tggatgtcg ttagtggcgc gcaatggcgc cgtgcgcgtc 4800
 ctgcgttgc tgccggatct tgaatgtcact ttcgtatgcg ttcttgcgt tgccggccat 4860
 gatatacgt tttggctgt ttagttgttgc tccatgttgc tggcccgatc tggtggatgc 4920
 ctcccttgcgtc tcatgtccatc tcaatgtcgat ggggttgcacc aggggtgcgc cctcgaaactt 4980
 cacctcggcg cgggttgcgtc agttgcgcgc gtccttgcag aatgtggcgc gtcctggac 5040
 gtagcctcg ggcatggcgg acttgaagaa gtcgtgcgtc ttcgtatgttgc cggggtagcgc 5100
 ggcgaagcac tgccggccgt agccgaagggt ggtcacgcgg gttggccagg gcacgggcag 5160
 ctggccgtg tgcaatgtt aatcggatgtt cagcttgcgg taggtggcat cgcctcgcc 5220
 ctgcggcggac acgtcaact tttggccgtt tacgtcgccg tccagctcgat ccaggatggg 5280
 caccaccccg gtgaacagct ctcgcctt gtcaccatc gttggcgatc cttgtatgttgc 5340

cagtggctgg gggttgcgcc gcccgggttt ataggaagcc acagcggcca ctcgagccat 5400
aaaaggcaac ttttaggaacg gcggggggtg attggattcg agtcgttta tcaccggcct 5460
tgccgcacag tgacgcattt ttttaccccc tcctccctcc ttttgcgggg gaaaaaaaaa 5520
aaaaaaaaaa aaaaaggaga agagaaaaaa agcgagcgag agagaaaagcg agattgagga 5580
agaggatgaa gagtttggcg atgggtgctg gtcccgtagg cccagatgga caagaatagc 5640
ccccgcctt gcggacagta tcccatcag tgactcagat cagataaagc ggccgcccagt 5700
gtgatggata tctgcagaat tccagcacac tggcggccgt tactagtgga tccgtacaag 5760
acagttagct agttggcaaa accaaggaga attctgtgac aaggaaaact caggacaaat 5820
cctgctaggc agttctggag gggaaagcgcc ctgcgcctgc agtactgacc tggaccata 5880
gccagggcag caatgcagga gaaggcagaa atgtcgatgt tcatattctg caagttggag 5940
gagaattcaa caatggaatc aatccattcc ccaaagccac gaacgcattg caacctgtgc 6000
aagaccaccc cattgcaaaa gatgagttt ccctccactg gttggacct gcaattaata 6060
ccaaagagag agaggggaga aaaagagaga gagaaaagct aaacaactaa ttacttgaag 6120
agcaataaat gagggattt aagtcaccc aattactgaa gagttatta aatgttagat 6180
ccagtggacc ttgaaagggt ttaatttcat aacaatcaag aacgccccct atcccacctc 6240
cattccattc atccaaggct tctggctcg ctcttctcg tcttttttgc ttttctttct 6300
tttttcttct tttecgact tecatttcct attctgtctt tttcttctacc ccaccctcg 6360
gtttcccttc ctcccttct ttccttct tgatttctct cacagcctcc ctggattgtc 6420
tcctccctcc cttattacct gtatgtaat cgaaggacaa acagttctaa gaaagctgt 6480
tcaaaaagca ggttttggc ggcttggc agtctgcga agccaggat cttctctgccc 6540
cagccccgga tgatctccat ggagccagtc aggagatcat agaattgctg gatatgtgg 6600
gtgtcatctc cactcattt gatgtcaggg ttcgcctgga actggattt cattttaaa 6660
agcacttaat gaggttctct aaaatatata acccgtgaa ttgctaacc cgtttctaat 6720
aggggagcca ggttttata acaattaaac ctctctctgaa ccagtaagga aattagatgt 6780
tccggggcaa ttaattcaat tagggatggt gctacgaggt cgctgctta aatatgtaaa 6840
ttgtcatttc catacagact aaatacagtg ccatccagcc ctggaaaacg ttcaactctt 6900
tctccacaaa acaatttgaca ttttagtatt catgctagtc ccagagtggg ctcgagctg 6960
agacgcctt ttgcaaatct taataaaatt gcagtcctta ccaggttgc tcaacatact 7020
tcctccaatt aaagcccgca tattttcac agtggagta aaaggatctg accacttgat 7080
cccccacccc actggcatct ttacaccctc tcctccctt ccctgacccc cagacagccc 7140
ccacatcctt ccccagcaca gtgcaggaa ggatgaggat taaagcatct atctctatgg 7200
aggactggga ataaaagttaga ttgggaaggg ggagagtttgc cacagccccct ggatttcatg 7260
gtagaggcac tggtgtatata ggaagaggag gggcgtatgg gtagtgggt agtggaaatt 7320
cacacagctt caaaccctgt ggcgtgtccc ttcctctt ctcctatct gtcctcatgt 7380
ctatgcffff agctacatgg atttgtctga agatctgaat cttgaggccc tggccagagc 7440
tgcgaggcat atacagccctt gttgcctt cttaaaaaaaa gttgaatctg agagttaatg 7500
acggatgtgg ggaggggtcc tgccatctg tcggttgtc cacatgatata cccccccgccc 7560
agcttcttac cttggatata gtcaggctgg tcatacgccgg gttggactcg acatgggccc 7620
tgacgaggc actgtatcaga ttcacccggg gcgaagggggg agagggtctc tggggctct 7680
tcggtttgc gggcaaaacga cctctccggc cttaaaact gtctgtgcga accactgcaa 7740
aggaagagcc ctgttagcgc cgttttccg agcccaggcc cagctgtgc ctgggtccct 7800
ccccggggaa ggccgcagcc gggggcacc aggtgtggc gctgaggggcc ccagtgttt 7860
taaaggcttc actgactaga atgggtctga tagttaaagg agagaaggc ctggggctt tctcttaggg aggccggggca 7980
agcaggcagc tgacgggtcc gggaggccat actgagggggg agtcggagat ccccgacacc 8040
ggaaagtgg acgtgtatgt ggatgttag ctttttttttgc ttaaaaggcgc aaactggagg 8100
gtcggcagct ccctcagcc taccttctt ggtacccggc gatacgtca actgtcttt 8160
acctttgtta ctactcttt tagaggcata tcaatctcca ctgaaggccaa tctatctgt acggcatctt tattcacatt 8220
atcttgcata aataatctgt ttaacaatgc ttttatatcc tgtaaagaat cattttcaa 8340
aatcatgtca aggtttctc gaggaaaaat cagtagaaat agctgttcca gtcttcttag 8400
ccttgatcc acttctgtca gatgtccct agtcagcgga gacctttgg ttttgggaga 8460
gtagcgacac tcccgaggat ttttcagaca ctggcgcac ttcgggtttt ctttggagca 8520
cttgagcttt ttaagtccgc aaatatcgca tgcttgcgtc atagaagaca gtagcttcat 8580
ggtggcaagc ttaagttaa acgctagaac tagtgatcc ggataagcca gtaaggcgt 8640
ggttctctag ttagccagag agctctgtt atatagaccc cccaccgtac acgcctaccc 8700
cccatggcg tcaatggggc ggagtttta cgacattttg gaaagcccg ttgattttgg 8760
tgccaaaaca aactcccatt gacgtcaatg gggtgagac ttggaaatcc cctgtgacta 8820

aaccgctatc cacgcccatt gatgtactgc caaaaccgca tcaccatggt aatagcgatg 8880
 actaataacgt agatgtactg ccaagttagga aagtccata aggtcatgt a ctgggcataa 8940
 tgccaggcgg gccatttacc gtcattgacg tcaatagggg gcgtacttgg catatgatac 9000
 acttgatgt a ctgccaagt ggcagttac cgtaaatact ccaccattg acgtcaatgg 9060
 aaagtcctta ttggcggtac tatgggaaca tacgtcatta ttgacgtcaa tgggcgggg 9120
 tcgttggcg gtcagccagg cggccattt accgtaagt atgtaacgcg gaactccata 9180
 tatggctat gaactaatga ccccgtaatt gattactatt aataactagt caataatcaa 9240
 tgtcaacccc gggctgcagg aattctaccg gtagggggag ggcgtttcc caaggcagtc 9300
 tggagcatgc gcttagcag ccccgctggc acttggcgc a tcacaagtgg cctctggcct 9360
 cgcacacatt ccacatccac cggtagcgcc aaccggctcc cttctttggt ggcccccttcg 9420
 cggcaccttc tactcctccc ctatcaggaa atgtcccccc cggcccgag ctcgcgtcgt 9480
 gcaggacgtg acaaatggaa gtagcaacgc tcaactagtct cgtgcagatg gacaagcacc 9540
 gctgagcaat ggaagcggtt aggcccttgg ggcagcggcc aatagcagct tggctccttc 9600
 gcttctggg ctcagaggct gggaaagggtt ggtccgggg gcccggctcag gggcgggctc 9660
 agggcgggg cggcgcgaa ggtcctccgg acccggcatt c 9701

<210> 74
<211> 2379
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic nucleotide sequence for GR region of homology

<400> 74
cgtacaagac agtagctag ttggcaaaaac caaggagaat tctgtgacaa gggaaactca 60
ggacaaatcc tgcattggcag ttctggaggg gaagcgccct ggcgcctgcag tactgacctg 120
tgaccatagc caggcagca atgcaggaga aggcaaaat gtcgatgttc atattctgca 180
agttggagga gaattcaaca atgaaatcaa tccattcccc aaagccacga acgcattgca 240
acctgtgcaa gaccacccca ttgcaaaaaga tgagttacc ctccactggg ttggacactgc 300
aattaatacc aaagagagag aggggagaaaa aagagagaga gaaaagctaa acaactaatt 360
acttgaagag caataaatga gggatttaag agtcaccta ttactgaaga gtaattaaa 420
atgtagatcc agtggacctt gaaagggtt aatttcataa caatcaagaa cgcggccat 480
cccacccca ttccattcat ccaagggttc tgggctegct tcttctcgct ttttttgg 540
ttctttcttt ttcttcctt ctccgacttc catttcctat tctgtcttt tctctacc 600
accctctggc ttcccttcctt ccctttctt tcctttctt atttctctca cagcctccct 660
ggattgtctc ctccctccct tattacctgt atgtaatcg aaggacaaac agttctaaga 720
aagctgattc aaaaagcagg tcttggtcgg ctttgggcag gtctgcaag ccaggatct 780
tctctgccc gccccggatg atctccatgg agccagtcg gagatcatag aattgctgga 840
tatgtgggt gtcatttcca ctcatttgat agtcagggtt cgcctggaac tggaaatttca 900
ttttaaaag cacttaatga ggttctctaa aatatataac ccgtgaaatt gctaacc 960
tttctaatag gggagccagg ttttataac aattaaaccc tctctctgacc agtaaggaaa 1020
ttagatgttc cggggcaatt aattcaatta gggatggtc tacgaggtcg ctgctttaaa 1080
tatgtaaatt gtcatttcca tacagactaa atacagtgcc atccagccct gggaaacgtt 1140
cactcttac tccacaaaaac aattgacatg ttagtattca tgctgtccc agagtgggccc 1200
tcgagctgag acggccctttt gcaaatctta ataaaatttc agtccctacc agtgggtctc 1260
aacatacttc ctccaaatcc a gccccgcata ttttcacag ctggagttaa agatctgac 1320
caattgtatcc cccacccac tggcatctt acacccttc ctcccttcc ctgacccccc 1380
gacagcccccc acatccttcc ccagcacagt gccaggaagg atgaggatta aagcatctat 1440
ctctatggag gactggaaat aaagtagatt ggaagggggg agagtttgc cagccctgg 1500
atttcatggt agaggcactg gtgtatgtt aagaggaggg gcgatggggat agtgggttag 1560
tggaaattca cacagcttca aaccctgtgg cgtgtccctc tcctcttcc tccatctgt 1620
cctcatgtct atgccccca gatcatggat ttgtctgaag atctgaatct tgaggccctg 1680
gccagagctg cgaggcatat acagccttgc ttgccttct ttacccctgt tgaatctgag 1740
agttaatgac ggtgtgggg aggggtctg cccatctgtc ggttgtcca catgatatcc 1800
ccccccggccag cttcttaccc tggaaatagtc caggctggtc atagccgggt tggagtcgac 1860

atgggcctg acgagggcac tgatcagact caccggggc gaaggggag agggctcctg 1920
 tgggcttgcagg gcaaacgacc tctccggcct tttaaactgt ctgtgcgaac 1980
 cactgcaaag gaagagccct gtagcggcg ctttccgag cccaggccc gctgctgcct 2040
 cggccctcc ccgggaaagg ccgcagccgc ggggcaccag gctgagccgc tgagggcccc 2100
 agtgctgt aagccttac tgactagaag cattaaaaaa tgccgggtta ttttatgtct 2160
 tcctccaaat gggtcgata gttaaaggag agaaggccct ggcggcttc tctagggaaag 2220
 gccgggcaag caggcagctg cagggtcctg gaggccatac tgagggggag tcggagatcc 2280
 ccagcaccgg gaagtggAAC gtgatgtgg agtatgagca gtggttcct aaaggcgc当地 2340
 actggagggt cgccagctcc ctcagccta cttctttt 2379

<210> 75
 <211> 9597
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 pCDGal4-DBD-GR sequence

<400> 75
 taacaggatt atttgtacaa gataatgtga ataaagatgc cgtcacagat agattggcTT 60
 cagtgagac tgatgcct ctaacattga gacagcatag aataagtgcg acatcatcat 120
 cggaaagagag tagtaacaaa ggtcaaaagac agttgactgt atcgcgggtt accattcagc 180
 aggccactac aggagtctca caagaaacct ctgaaaatcc tggtaacaaa acaatagttc 240
 ctgcaacgtt accacaactc acccctaccc tgggtcact gttggagggtt attgaacctg 300
 aagtgtata tgaggatgat gatagctctg ttccagactc aacttggagg atcatgacta 360
 cgctcaacat gttaggaggg cggcaagtga ttgcagcagt gaaatggca aaggcaatac 420
 cagtaagat gcaaaacata aaagagacaa tatataaaacc tttgtgtttt cttcagcaaa 480
 aacacttgg cttttatatac atcgtgagcc catggcttat cttgtttctc tttagttctgg 540
 ggactatgaa ggggagagtc aggtgaatac agtgatagg gagtttataa taaaacattt 600
 acattactcc ctgctttca aatcattatg cacaggatgg taatttcaca taggatgatg 660
 taatatcaga attcaagttt caagactcac tcaaaactcc ttttacactg aagtttgggg 720
 aaagaaaatg ttttagtta attccatttg ttttccttca ttgtgccact tttaaaaatc 780
 aggttggggtaaagattttaa acatcaag tatgttgcattt gtcaaaatattt gtactaaagt 840
 agaatgattt taacccttca ctaaatgaaa tgctacacat tgaatgtat tttaaaagata 900
 attttaataa aaagttaccc tattggatt tgggtgtggaa tggcagaggtt caatgttagt 960
 gtcagctctg actttaaaga caggaaattt gcaagcctgt gttcacgcaa atagtttaggg 1020
 agagagaacaaatgtttaa acatcaag ttttgcatttgc ttttgcatttgc ttttgcatttgc 1080
 gtgaatagca gggcaaatgtt tttgcatttgc ttttgcatttgc ttttgcatttgc ttttgcatttgc 1140
 ttttgcatttgc ttttgcatttgc ttttgcatttgc ttttgcatttgc ttttgcatttgc ttttgcatttgc 1200
 taggggtac atgtatataat ttaataaaag catacaacgt gtaaggatca aatcagatgt 1260
 actgggatat ccatcacctc aaacattgtt ttggggaca ttccaaatct tctcttttag 1320
 ctatattgaa atataaagta aattattgtt aactatagtc atcctgttgc tctactgaac 1380
 actaaaactt attcttctca actgtatattt tgcacccctgc aaccattccc gttcatccc 1440
 catcaccact atcttcccg gtcactgtt accgccaagc caagaattttt ggctattttt 1500
 ctatattgtt catgttact taagcagaca gaggtgacaa aactggctttt tttttttttt 1560
 ttttacatta aaagctttaaaaaccccttggggctgg gtcgcatttgc tcaagcctgt 1620
 aatcccacca ctttggggaaag cccagggtggg tggatcgtt gaggtcagga gttcgagacc 1680
 agccctggcca gcatagcaaa accccatctc tactaaaattt aaaaaattt gcccggcatg 1740
 gtggatgaa tctgttattcc tagctacttg ggaggctggc actgagaatc acttgaaccc 1800
 gggaggcgga ggttgcagt gacccgatgc gcaccattgc actccagcctt ggggtgacaga 1860
 gcaagactttt gtctcaattt aaaaaaaaaaaaaaaa aacacaagag ggtttgcgt 1920
 tcttaaagtgtt tcaatgttca gaaatgttactt gttgttgcatttgc ttttgcatttgc 1980
 tggtaggggtt ggccttgcatttgc ttttgcatttgc ttttgcatttgc ttttgcatttgc 2040
 acagagatgtt tcttctgttca gttgttgcatttgc ttttgcatttgc ttttgcatttgc 2100
 aggaacttac acctggatgttca gcaatgttactt gttgttgcatttgc ttttgcatttgc 2160
 gcatattgttca gttgttgcatttgc ttttgcatttgc ttttgcatttgc ttttgcatttgc 2220

gctccggcg tatatgctcc gcattggtct tgaccaactc tatcagagct tggttgacgg 5760
 caatttcgtat gatgcagctt gggcgcaggg tcgatgcgac gcaatcgtcc gatccggagc 5820
 cgggactgtc gggcgtacac aaatcgccc cagaagcgcg gccgctgga ccgatggctg 5880
 tgtagaagta ctgcgcgata gtggaaaccg acgccccgc actcgccga gggcaaaggaa 5940
 atagcacgtg ctacgagatt tcgattccac cgccgccttc tatgaaaggt tgggcttcgg 6000
 aatcgtttc cgggacgcg gctggatgat cctccagcgc ggggatctca tgctggagtt 6060
 cttegcaccc cccaacttgt ttattgcage ttataatggt tacaataaa gcaatagcat 6120
 cacaatattc acaaataaaag catttttc actgcattc agttgtggg tgcattaaact 6180
 catcaatgtt tcttatacc tctgtatc gtcgacactc agcttagact tggcgtatc 6240
 atggtcatacg ctgtttctg tttgaaattt ttatccgctc acaatccac acaacatacg 6300
 agccgaagc ataaaagtta aagcctgggg tgccataatga gtgagctaac tcacattaat 6360
 tgcgttgcgc tcaactgccc ctttccagtc ggaaacctg tcgtgccagc tgcattaaat 6420
 aatcgccaa cgcgcgggaa gaggcggtt gcttattggg cgcttcccg cttcctcgct 6480
 cactgactcg ctgcgtcgg tcggttcggc gcccgcagcg gttatcagtc actcaaaggc 6540
 ggttaatacgg ttatccacag aatcagggaa taacgcagga aagaacatgt gagaaaagg 6600
 ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgcgt gcttcccttcc ataggctccg 6660
 cccccctgac gagcatcaca aaaatcgacg ctcaagtca gaggcgaa acccgacagg 6720
 actataaaga taccaggcgt tccccctgg aagctccctc gtgcgtctc ctgttccgac 6780
 cctgccctt accggatacc tttccgcctt tcccccctcg ggaagcgtgg cgcttctca 6840
 tagctcacgc tttaggtatc tcagttcggt tttaggtcggt cgctccaagc tgggctgtgt 6900
 gcacgaaccc cccgttcagc ccgaccgcgtg cccttatcc ggttaactatc gtcttgagtc 6960
 caacccgtta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag 7020
 agcgaggtat tttaggcgtt ctacagagtt ctgttgttgg tggcctaact acggctacac 7080
 tagaagaaca gtatttggta tctgcgtct gctgaagcca gttacccctcg gaaaaagagt 7140
 tggtagctct tgatccggca aacaaaccac cgctggtagc ggtttttttgg ttgcaagca 7200
 gcagattacg cgccggaaaa aaggatctca agaagatctt ttgatctttt ctacggggtc 7260
 tgacgctcag tggAACGAAA actcacgatc agggattttt gtcatgagat tatcaaaaag 7320
 gatcttacc tagatccctt taaattaaaa atgaagttt aatcaatct aaatatatata 7380
 tgagtaact ttgtctgaca gtttccatc ttaatcagt gagggccacta ttcagcgat 7440
 ctgtcttattt cgttcatcca tagttccgtt actcccccgtc gtgttagataa ctacgatacg 7500
 ggagggttta ccattctggcc ccagtgtcgtc aatgataccg cgagacccac gtcaccggc 7560
 tccagattta tcagcaataa accagccagc cggaaaggggc gagcgcagaa gtggctctgc 7620
 aactttatcc gcctccatcc agtctattaa ttgttgcgg gaagcttagag taagtagttc 7680
 gccagttaat agtttgcga acgttgcgtc cattgttaca ggcattgtgg tgcacgctc 7740
 gtcgttttgtt atggcttcat tcagctccgg tttcccaacga tcaaggcgag ttacatgatc 7800
 ccccatgttg tgcaaaaaag cggttagtctt cttccgttcc cgcattgttgc tcaagatcaa 7860
 gttggccgca gtgttatcac tcattgttgc ggcagcactg cataattctc ttactgtcat 7920
 gccatccgta agatcgaaaa ctgtgactgg tgagttactca accaagtcat tctgagaata 7980
 gtgtatgcgg cgaccgagtt gcttttgcgg ggcgtcaata cgggataata cccgcacca 8040
 tagcagaact ttaaaagtgc tcatttttttgg aaaaacgttct tcggggcgaa aactctcaag 8100
 gatcttaccg ctgttgagat ccagtgcgtat gtaaaccact cgtgcaccca actgtatctt 8160
 agcatctttt actttcacca gcttttgcgg tgagggcaaaa acaggaaaggc aaaaatgcgc 8220
 aaaaaaggaa ataagggcga cacggaaatg ttgaaatactc atactcttcc tttttcaata 8280
 ttattgaagc atttatcagg ttattgtct catgagcgga tacatattt aatgtattta 8340
 gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgcac ctgacgtcga 8400
 cggatccggaa gatctcccgaa tcccttatgg tgactctca gtaaatctg ctctgatgcc 8460
 gcatagttaa gccagttatct gctccctgtt gttgtgttgg aggtcgctga gtgtgcgcg 8520
 agcaaaaattt aagctacaac aaggcaagggc ttgacccgaca attgcatgaa gaatctgttt 8580
 agggtttaggc gttttgcgtt gtttgcgtat gtaacggccca gatatacgcg ttgacattga 8640
 ttattgacta ttatataataa gtaataattt acggggtcat tagttcatag cccatatatg 8700
 gagttccgcg ttacataact tacggtaat gggccctgt gctgaccgc caacgacccc 8760
 cggccattga cgtcaataat gacgtatgtt cccatagtaa cgccaaatagg gactttccat 8820
 tgacgtcaat ggggtggagta ttacggtaa actgcccact tggcagtaca tcaagtgtat 8880
 catatgccaa gtaacggcccc tattgacgtc aatgacggta aatggccgc ctggcattat 8940
 gcccagttaca tgaccttatg ggactttccctt acttggcagt acatctacgt attagtcatc 9000
 gctattacca ttgtgtatgcg gttttggcag tacatcaatg ggcgtggata ggggtttgac 9060
 tcacggggat ttccaagtctt ccacccattt gacgtcaatg ggagttgtt ttggcaccac 9120
 aatcaacggg actttccaaa atgtcgtaac aactccgcacc cattgacgc aatggccgg 9180

aggcgtgtac ggtgggaggt ctatataagc agagctctct ggctaactag agaaccact 9240
 gcttaactggc ttatcgaaat taatacgact cactataggg agacccaagc tggctagcgt 9300
 ttaaacttaa gcttgcacc atgaagctac tgtcttctat cgaacaagca tgcgatattt 9360
 gccgacttaa aaagctcaag tgctccaaag aaaaaccgaa gtgcgccaag tgtctgaaga 9420
 acaactggg a gtgtcgctac tctccaaaaa cccaaaggtc tccgctgact agggcacatc 9480
 tgacagaagt ggaatcaagg cttagaaagac tggAACAGCT atttctactg atttttcctc 9540
 gagaagacct tgacatgatt ttgaaaatgg attctttaca ggatataaaa gcattgt 9597

<210> 76
<211> 10872
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
pKI-Gal4-DBD-GR sequence

<220>
<221> modified_base
<222> (8885)..(8885)
<223> any nucleotide

<400> 76
cagttcgatg taacccactc gtgcacccaa ctgatcttca gcatcttta ctttcaccag 60
cgtttctggg ttagaaaaaa caggaaggca aatgccgca aaaaaggaa taagggcgac 120
acggaaatgt tgaatactca tactcttct tttcaatat tattgaagca tttatcaggg 180
ttattgtctc atgagcggat acatatttg aatgtatgg aaaaataaac aaatagggt 240
tccgcgcaca tttccccgaa aagtgcocacc taaatttgtaa gcgtaataat tttgttaaaa 300
ttcgcgttaa attttgtta aatcagtc a ttttaacc aataggccgaa aatcgccaaa 360
atcccattata aatcaaaaaga atagaccgag atagggttga gtgtgttcc agtttggAAC 420
aagagtccac tattaaagaa cgtggactcc aacgtcaaag ggcggaaaaac cgtctatcag 480
ggcgatggcc cactacgtga accatcaccc taatcaagtt ttttgggtc gaggtgcgcgt 540
aaagactaa atcggaaccc taaaggagc ccccgattta gagcttgacg gggaaagccg 600
gcgaacgtgg cgagaaaagga agggaaagaa gcggaaaggag cgggcgcgtag ggcgcgtggca 660
agttagcgg tcacgctgog cgttaaccacc acacccggcc cgcttaatgc ggcgcgtacag 720
ggcgcgtccc attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggccc 780
tcttcgttat tacgcccagct ggcgaaagggg ggtatgtctg caaggggatt aagttgggt 840
acgccagggt tttcccagtc acgacgttgc aaaaacgcgg ccagttagcgc cgcgtataac 900
gactcactat agggcgaatt ggagctccac cgcgggtgcgg ccggccatg caggccacga 960
catgataaga tacattgtat agtttggaca aaccacaact agaatgcagt gaaaaaaaaatg 1020
ctttattttgt gaaattttgtg atgctattgc ttattttgttta accattataa gctgcaataa 1080
acaagtctt gtacagctcg tccatgccga gagtgtatccc ggcggcggtc acgaactccca 1140
gcaggaccat gtgatcgccgc ttctcggtgg ggtctttgtc cagggcggtc tggtagctca 1200
ggtagtgggt gtcgggcgcg acgacggggc cgtcgccgtat ggggggtgttc tgctggtagt 1260
ggtcggcgag ctgcacgtg ccgtccctcgat ttttggcg gatcttgcgat ttcacccgt 1320
tgccgttctt ctgtttgtcg gccatgat agacgttgcg gctgttgcgat ttgtactccca 1380
gcttgcgtccc caggatgttgc cctgtccctt tgaagtcgtat gcccctcagc tcgtgcgggt 1440
tcaccagggt gtcgcctcg aacttcaccc cggcgccgggtt cttgttagtgc ccgtcgccct 1500
tgaagaagat ggtgcgtcc tggacgtacg cttcgccat ggcggacttg aagaagtgcgt 1560
gctgcttcat gtggcggtt tagcgggcga agcaactgcg gccgtagccg aaggtggtca 1620
cgagggtggg ccaggggcgcg ggcagcttc cgtgtggtgcgat gatgaacttc agggtcagct 1680
tgccgttaggt ggcacatcgccc tcgcctcgat cgcacacgcgt gaaacttgcgat ccgtttacgt 1740
cgccgtccag ctcgaccagg atgggaccca ccccggtgaa cagctccgtcc cccctgctca 1800
ccatgggtggc gttagcttgc tttgacgttgc gctgggggtt ggcggccgg gttttatagg 1860
aagccacagc ggcacactcgat gccataaaaag gcaacttttag gaaacggcggtt ggggtgattgg 1920
attcgagtcg tttatttcacc ggccttgcgc cacagtgcgcg catttttta ccccccctctcc 1980
cctcccttttgc cggggaaaaaa aaaaaaaaaaa aaaaaaaaaaa ggagaagaga aaaaaaagcga 2040

gcgagagaga aagcgagatt gaggaagagg atgaagagtt tggcgatggg tgctgggtcc 2100
gttaggcccg atggacaaga atagcccccg cccttgcgga cagtatccc ttcaagtact 2160
cagatcagat caagcggccg cagggcagcag catgagaaca caagaatagg aagtacatat 2220
cccttaatga aaatgaaact aatgtaaatc tggtgaggaa aaaaaattat aaatcccact 2280
ttatgttaagg aatgtatgtg gcatgattag attttgtgt gactgctta agctaaatct 2340
gaatgccta ataataaaaa ccctaaatta gtttagttt ataaaaccagg ataagtttgt 2400
gtttcataca aactgttta tgctttaat tcacactaaa tgccttcaa ttacatatta 2460
gaaagtgact aaaaaacaaa tactgccaac atggtatttt ttaaacagca tggttaggatt 2520
atgatgaga ataatgtctt cagttcccc agaaaactaaa tatatttacc agttaattt 2580
aaattggatc caactaccta atcttggcac attgtgatgc agtagttga acagtcctct 2640
tttggcaat aaaagtcaag gtgacttctg tttctatac aaacagaaat ctcccttcatt 2700
tttatttctt cttagaccag gggtagtgt gtgtgaactt cctgtttaca ttgtactaat 2760
agcaataaagc taggggtacg aaaagttagt atacaaggta gaaagtgaga gtaacttagga 2820
tgtgacttct ctgctgtgt aaggatggat gggaatgaag tctggatg tgaacaccaa gagtctaaga gtatgctgga 2880
attctggaaa gtcgaaatc agacaaatctg tgcactgggg agatgtgaa gggaaaacaat ggagaggaa gtgaggctcg 2940
attctggaaa gtcgaaatc agacaaatctg tgcactgggg agatgtgaa gggaaaacaat ggagaggaa gtgaggctcg 3000
attagccaaatc atgatgaaacg ttacactgga gggagaaaaac gagtgaggaa acaactgcaa 3120
tgaggtaagg gggaaatggac acctgaagta tttcgacac agacaaataa gcttagaggaa 3180
taagagaata gctgatttat aaaatcttaa agcttggtc ttctctcagt taaacaaaac 3240
aaatttggtc ccatattctt aaaccacaga cactgggtga gccttgccat gcaaagagat 3300
acctcttctg tctgggtcca tatgtcccac cccatcagcc tagatatcct catctacaat 3360
tgtcccaact gaatccacat gagtttacc atgctcctt ttcttctgg aatagggtta 3420
tgtaaatgtat atoacagtaa gattcctgt tgcttgcgac agacttact aacctcagat 3480
atgctgttaa atatcctcaa acatgtctg ggacagtaat aatgccacat gtacatttgc 3540
tttacattac acttgcctg ccccaagcac tcataactct atttcagtt tgcttagatc 3600
ctagataacct agtaggattt ttcagtcct gaattatcac aatagcagtc aatcaggaaa 3660
acatcagctg gttaaagata ccctatgaat acagggaaaa tgacacacat acaacttact 3720
cattaataat cagatcagga gcaaaacaca gcaggtttgc acttgattgt ctatatgatc 3780
tccaccccg agcaaaatgcc catccaggtg taagttcctg ataagaaaaca tccaggagta ctgcagtagg gtcatttgg 3840
ttcagaagat catctctgtg aaacctgaat taagagaat aaaggatgaa ggcacacactc 3900
aacaaggcga cccctaacacg ggaattgcca aggagcaatg agatcaatta gccctgtcaa 3960
gtcatctgac actttaagac aaaatggaaa ttaaataacta ggtagacaca gttttctt 4020
aattgagaca aagtcttgcg tgcacccca ggctggagtg caatggtgc atctcggtc 4140
actgcacaccc cccctcccg ggttcaagtg attctcagtg ccagccccc aagtagctag 4200
gaatacagat tcataccacc atgcccggct aatttttgc attttagtag agatgggtt 4260
ttgctatgt ggcaggctg gtctcgaact cctgacccca actgatccac ccacctggc 4320
ttcccaaagt gctggatta caggcgtgag ccatcgcacc cagcccccta ggtcttttt 4380
aatagcttt aatgtaaaaaa aaaaaaaaaa aagccagtt tgcacccctt gtctgcttaa 4440
gttaacatga actaaatagt aaaatagcca aaattcttg cttggcggtt accagtgacc 4500
ggaaagata gtgtgtatgg agaagaaaata agtttagtgc ttcagtagca caacaggatg actatagttt acaataattt 4620
actttatatt tcaaaaatagc aggtgatgg tatcccagtt taaaagagaa gatttggat gttccccaaa caaatgttt 4680
tattatcatg taccctctaa actctgattt gatccctaca cttttttttt tttttttttt 4740
tgattnaaaa aaacagaaaaa atataaataat gtacagtaaa tatgtacaac tattatgtat 4800
catttgcct gctattcaca gatcttgcgaa aatcaatggatg tgcgttgc aatggatgg 4860
ggttacttt ttgtctctc cttaaatgtc agagctgaca ctaacattga cctctgccc tccacaccaa attccaaatag 5040
ggtaactttt attaaaaattt atctttaaaa ttacattca tgcgttgc ttcattttgt 5100
gaagggttaa aatcattcta ctttagtaca aattttgaca atcaacatac ttgtatgttta 5160
ccaatcttac aaacaacctg attttaaaa gtggcacaat gaaggaaaac aatggaaatt 5220
aactaaaaac atttcttcc aaaaaacttc agtgtaaaag gagttttag gatcttgc tttttttttt 5280
aacttgaatt ctgatattac aatgttttat tataaactcc cttttttttt tttttttttt 5340
gaaaagcagg gagaatgtt actctccct tcatagtccc cagaactaag agaaaacaaga taagccatgg gctcacgt 5400
atataaaaagc caaagtgttt ttgctgaaga aacacaaaag gtttatatacg ttgctttttt 5460
ttgctgtttt

atgtttgca tcttacctgg tattgcctt gcccattca ctgctgaat cacttgcgc 5580
cctcctaaca tggtagcgt agtcatgatc ctccaaggta agtctgaaac agagctatca 5640
tatccctgcat ataacacttc aggttcaata acctccaaca gtgacaccag ggttaggggt 5700
agttgtgta acgttgcagg aactattgtt ttgttaccag gattttcaga ggtttctgt 5760
gagactcctg tagtggcctg ctgaatggta cccggcgata cagtcaactg tctttgacct 5820
ttgttactac tctttccgta tgatgtgtc gcacttattc tatgtctgtt caatgttaga 5880
ggcatatcag tctccactga agccaatcta tctgtgacgg catctttt cacattatct 5940
tgtacaaaata atcctgttaa caatgtttt atatcctgtt aagaatccat ttcaaaatc 6000
atgtcaaggt ctctcgagg aaaaatcagt agaaaatagct gttccagtct ttctagcct 6060
gattccactt ctgtcagatg tgccctagtc agcggagacc ttttggtttt gggagagtag 6120
cgacactccc agttgttctt cagacacttg ggcacttcg gttttctt ggagcacttg 6180
agtttttaa gtcggcaaatt atcgcacgt tttcgatag aagacagtag cttcatgggt 6240
gcaagcttaa gtttaaacgc tagaactagt ggtatccggat aagccagtaa gcagtgggtt 6300
ctctagttag ccagagagct ctgcttatat agacctccca ccgtacacgc ctaccgcccc 6360
tttgcgtcaa tggggcggag ttgttacgac attttggaaa gtcgggttga ttttgggtcc 6420
aaaacaaact cccattgacg tcaatgggt ggagacttg aaatccccgt gagtcaaacc 6480
gctatccacg cccattgatg tactgcacaa accgcacac catggtaata gcgtatgacta 6540
atacgttagat gtactgccaat gtagggaaatg cccataaggt catgtactgg gcataatgcc 6600
aggcgggcca tttaccgtca ttgacgtcaa tagggggcggt acttggcata tgatacactt 6660
gatgtactgc caagtgggca gtttaccgtt aataactccac ccattgacgt caatggaaag 6720
tccctattgg cgttactatg ggaacatacg tcattattga cgtcaatggg cgggggtcgt 6780
tgggcgtca gccaggcggg ccatttaccg taagttatgt aacgcggaaac tccatatatg 6840
ggctatgaac taatgacccc gtaatttgatt actattaata actagtcata aatcaatgtc 6900
aaccggcggc tgcaaggatt ctaccggta ggggaggcgc ttttcccaag gcagtctgg 6960
gcatgcgtt tagcagcccc gctggcaactt ggcgcacac aagtggcctc tggcctcgca 7020
cacattccac atccaccggc agcgccaaacc ggctcccttc tttgggtggcc ctttcggcc 7080
accttctact cctcccttag tcaggaagt ccccccggcc ccgcagctcg cgtcgtgcag 7140
gacgtgacaa atggaagtag cacgtctcac tagtctgtt cagatggaca agcaccgctg 7200
agcaatggaa ggggttaggc ctttggggca ggggcaataa gcagcttgc tccctcgctt 7260
tctgggtca gaggctggga aggggtgggt cccggggcgg gtcaggggc gggctcagg 7320
gcggggcggg cgccaaaggc tcctggaccc ggcattctgc acgttcaaa agcgcacgtc 7380
tgcgcgttgc ttctctt cctcatctt cggccttcgaa cctgcataaa aaagcctgaa 7440
ctcaccgcga cgtctgtca gaagtttctg atcgaaaatg tcgacagcgt ctccgacactg 7500
atgcagctt cggaggcga agaatctcg tcttcagct tcgatgttagg agggcgttgg 7560
tatgtcttcg gggtaaatag ctgcgcgtt gtttctaca aagatcgta tgtttatcg 7620
caacttgcata cgccgcgtt cccgattccg gaagtgtttt acattggga attcagcgg 7680
agcctgaccc attgcatactc ccgcgttgcga cagggtgtca cgttgcataa cctgcctgaa 7740
accgaactgc ccgcgttgc gcagccggc gggaggcga tggatgcgt cgctgcggcc 7800
gatcttagcc agacgagcgg gttcgccca ttccggaccgc aaggaatcg tcaatacact 7860
acatggcgtt atttcatatg cgcgattgtt gatccccatg tgtatactg gcaaaactgtg 7920
atggacgaca ccgtcagtgc gtccgtcgcg caggctctcg atgagctgtt gctttggcc 7980
gaggactgcc ccgaagtccg gcacccgtt caccggatt tcggctccaa caatgtcctg 8040
acggacaatg gcccataac agcggttattt gactggagcg aggcatgtt cggggattcc 8100
caatacgggg tgcacaatctt cttttctgg agggcgttgg tggcttgtat ggagcagcag 8160
acgcgttact tgcaggcggag gcatccggag ctgcaggat ccgcgcggc cccggcgtat 8220
atgcgtccgca ttggcttgc ccaactctat cagagcttgg ttgacggcaat ttgcgtat 8280
gcagcttggg cgccagggtcg atgcacgcgat atcgatccgtt ccggagccgg gactgtcggt 8340
cgtagacaaa tgcggccgtt aagcgccggc gtcgtggaccgc atggctgtt agaagacttc 8400
ggcgatagt gaaaccggcgc ccccgacact cgtccggaggca caaaggaaatg cctggagaaag 8460
gaagtgttgcgtt gtaaaaggctg agctctctt ctgcgtatg tagcctctt ttagctcg 8520
caacttgcgtt ttttgcgtt cgcgttgcgtt ccgcgttgcgtt agtctgcata 8580
aattgtatgtt ctattaaaca ataaagatgtt ccactaaaaat ggaagttttt tactgtcata 8640
ctttgtttaag aagggtgaga acagagtttcc tacatgttgc atggaggat tggagctacg 8700
gggggtgggg tgggggtggg attagataaa tgccgtctt ttactgttgcgtt ctctttacta 8760
ttgcgttgcgtt ataatgtttc atagttggat atcataattt aaacaagcaa aaccaaaat 8820
agggccagctt catttcgttca ctcgtatgtt atagatctt atgcgttgcgtt tggatcatt 8880
gtttntctgaa tccactggaa gcttgcgtt acgcgttgcgtt tcggagggggg gcccgggtacc 8940
cagctttgtt ccccttgcgtt gagggttaat tgccgttgcgtt ggcgtatcat ggtcatagct 9000

gtttcctgtg taaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 9060
 aaagtaaa gcctgggtg cctaattgat gagctaactc acattaattt cggtgcgtc 9120
 actgcccgt ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg 9180
 cgcggggaga ggcgggttgc gtattggcg ctcttccgtc tcctcgctca ctgactcgct 9240
 gcgctcggtc gttcggtgc ggcgaggcgat attagctcac tcaaaggcgtaataacggtt 9300
 atccacagaa tcagggata acgcaggaa gaacatgtga gcaaaaggcc agcaaaaggc 9360
 caggaaccgt aaaaggcccg ctgtgtggc gttttccat aggctccgccc cccctgacga 9420
 gcatcacaaa aatcgacgt caagtcagag ggccggaaac ccgcacaggac tataaagata 9480
 ccaggcggtt ccccttgaa gctccctcggtc ggcgttcctt gttccgaccc tgccgcttac 9540
 cggtacatcg tccgccttc tcccttcggg aagcggtggc ctttcata gtcacgctg 9600
 taggtatctc agttcggtgt aggtcggtcg ctccaaagctg ggctgtgtgc acgaaccccc 9660
 cggtcagccc gaccgctgog ctttatccgg taactatcg tttgagtcca accccgtaag 9720
 acacgactta tcgcccactgg cagcagccac tgtaaacagg attagcagag cgaggtatgt 9780
 aggccgtgct acagagttct tgaagtgggt gcttaactac ggctacacta gaaggacagt 9840
 atttggtatac tgcgtctgc tgaagccagt taccttcggg aaaagagttt gtagctttt 9900
 atccggcaaa caaaccaccc ctggtagccg tggtttttt gtttgcacg acgagattac 9960
 ggcgagaaaa aaaggatctc aagaagatcc ttgtatctt tctacgggggt ctgacgctca 10020
 gtggaaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac 10080
 ctagatctt taaaattaaa aatgaagttt taaatcaatc taaagtatata atgagtaaac 10140
 ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 10200
 tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac gggagggctt 10260
 accatctggc cccagtgtcg caatgatacc gcgagaccca cgctcaccgg ctccagatt 10320
 atcagcaata aaccagccag ccggaaaggc cgagcgcaga agtggctctg caactttatc 10380
 cgcctccatc cagtcttta attgttccgg ggaagctaga gtaagtagtt cgccagttaa 10440
 tagtttgcgc aacgttggc ccattgctac aggcacatcggt gtgtcacgct cgtcggttgg 10500
 tatggcttca tttagtcccg gttcccaacg atcaaggcga gttacatgtat ccccatgtt 10560
 gtgcaaaaaa gcggttagct ctttcgggtcc tccgatcggt gtcagaagta agttggccgc 10620
 agtgttatca ctcatggta tggcagcaat gataattctt ctactgtca tgccatccgt 10680
 aagatgttt tctgtgactc gtgagactc aaccaagtc ttctgagaat agtgtatgcg 10740
 gcgaccgagt tgccttgcc cggcgtcaat acgggataat accgcgccac atagcagaac 10800
 tttaaaatgt ctcatcattt gaaaacgttc ttccggggcga aaactctcaa ggatcttacc 10860
 gctgttggata tc 10872

<210> 77
 <211> 3250
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic nucleotide sequence for MR region of homology sequence

<400> 77
 tttgtggtgc taaaaaaatg agcatgtggt gcctcctatg gtccaggta ataccagact 60
 ttgataaagt tctatggtag aacatgtttc cctggcacct gtcctccgt cctgacagac 120
 tcaagctttt cttttgttct ttatttttt tattttttt ttaagatgaa ttctcatttt 180
 gtcgcccagg ctggagttca gtgacatgtatggctcac tgcaccccgat gttcaagcg 240
 attctcctgc ctcaggctca caagtagctg gggttacagg tgcctcccg gacgcccac 300
 taaaattttt attttttagta gagacagggt ttcatgtatgg tggcaaggct ggtctcaa 360
 tcctgaccc tcctgtttt aagtgtatcg cttgcctcg cctcccaaag tctgggatgg caggcatggg 420
 ccactgcacc aggccttctt ttgtttttt tcaacctcccc catctccagg ggcacatgt 480
 tatttttttgc cctgtggagt ttcaactggc aagctccctt tcaaagaatc agtcgttgg 540
 gctttccccca acttttacag ctgctgtctg accacaggta atctaagaaa ctgctttt 600
 ttaatgaact gggacaatgt aatggcaatt tctaagtgtat gtttctgttag attaagaggg 660
 gccttattaa taaaatcca taaaacaaaaaaaatgact aaaataggaa ttccaggagac 720
 aatttaaagg gtgctggggg ccaggtgatt attttttaga cagaagagtc attccaggtc 780
 tcttacaact tactgttagt gccttgcctt acttggtcta cacttagatt aaccctcggt 840

atgagccgct aataatgtgc tcaacttgc atttcatgaa cagaatgtat tttaccttat 900
 taatattttt gnatggaaatt ttccatattt atacaacataa gaagtgtatg tcattttttt 960
 tatttttaat ttctttttt ttttgagac agagtctcac tctgttgctc aggctggagt 1020
 atagtggtgg gatctcgct cactgcagcc tctgcccgc aggttcaagc gattctcctg 1080
 cctcaggctc ctgagtagct gggattacag gcacatgcca ccaagcagga ctaattttt 1140
 tgtttttagt agaaacgggg ttccaccatc ttggcttaggc tggcttgaa ctcctgacct 1200
 cgtgatccat ctgcctcagt aagccacctg cttggcctc ccaaagtgtc gggattacag 1260
 gctgagcca ctgcggccag ccgtatctgt cattttaaaca ataaaatctt taaccctgca 1320
 ttctcgaaag aaatgagcac attaaccata ttttcctat acacttgatc aatttggat 1380
 caatttaaaa aatctgatat caggattca tagaacgca actcctcctg catgggttgg 1440
 gatggcgaag tcagttgcc agacttggta agttcaggat gcagcctgtg aaaggagagg 1500
 caatcctacc tggaaagtacc tttgcccact tcacgacttg gatcatctgt ttgcctgcta 1560
 agcggtttag gctggagagc agattttcg ctgtatctgg ttttgagctg tcatagcctg 1620
 catatacaat ttcaaggatca atgtttcaaa ggaccataac gggggaaaggt gtgagcgctc 1680
 gtgagattgtt ggagagctga ggaaccatgt ctgtgttgcg cgagggttct ttgcaggag 1740
 cgatgtacgt tggcccttcc tctgggctt gccccgggtgg ggggtgggggt gggggctgct 1800
 gctgctgtgg ctgctccctg tgaatccctt tctttcattc ttttgacatc cttttcttgc 1860
 tttcatcacc aacctatttc tcatctgcct cttcccttac tggaaaccctt gttttcttat 1920
 tgccagagag accatctgt cagaatatgg ctgttagttac tatacataatt caaatattcc 1980
 ctttaagta tttcttgacc caacaattac aaagaagcac tgctagatac aaacataaaat 2040
 aaaatgtgct aatgaggaat gcaaaacagta taaaatgtaa aactttaaaat aggaaggaat 2100
 ttgtggccaa agaaaaactg attacgattc ttcaagtctt taaaatgca tcataaaaact 2160
 aaaaatgttt tgctacagaa taacagttt gcttctttt tggatggcaaa gaagggttaat 2220
 tacgatgtc atgaaaacac attgagcagc tggataactgt taacaaaaga accattataat 2280
 gagctcgtaa aagctggcac tttatataat cacatttccctt agaaaattaga ttgtcttccc 2340
 aattggagtc gataccaaaa gagactgttt tctttaattt gtggctttgtt accagtcaat 2400
 tttcagagcc aactggatca tctttctttt gggatataatccatcataat ttagtggaa 2460
 tttatataat ctgattttata tacactcaga agcatacagt ataatgtgc aaagagtaat 2520
 ttttaagtt cccagagatc tggtaatttt taacgtttaaa aatggccagt ttctgtat 2580
 tctttcacca acgtacatct gtttagaaat tactgaaaga aatcataacg cataactctg 2640
 cagtttattt tatgaaggct aattaatagg tacttactca taaaagacta ggtctgggtgc 2700
 aaaaatagaga aattggctgt tcgtatgtt gtacgatctc cagctcaagg caaatgtga 2760
 tagacacatc caagaataact ggattagggt aatttggtcc tcaagaggca agttttaaa 2820
 tcctgaagaa caaaaacaaatt aatcacaagaa atacacttag catttaagta cattccagga 2880
 agatgcttct taagtcaacc ccaaaacagcc acctttctt cactttctc ttttctgaca 2940
 gactcaaatc aattcaacag tcatttacag agtggccacc actggccctg tcctgttcca 3000
 catttcactt taaattcatc ctaaattcta gatgagggtt aactgcccagg gcaaaactaaa 3060
 attgcaaaatg tcagcaagtt tcagtgtgga tgacagccta aaacccttgg ttagattgag 3120
 agaatttggta ttaacctgtt aaacttgcctt cgtcccaccc acaattaata acactaaaat 3180
 aaagggttaat ttaataactg tcaaaccctg tggatctgca ataatggtgg gagatttcat 3240
 tggtcatgag 3250

<210> 78
 <211> 9212
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 pCDGal4-DBD-MR sequence

<400> 78
 agcactgcat aattctctta ctgtcatgcc atccgttggaa tgctttctg tgactgggtga 60
 gtactcaacc aagtcttctt gagaatagtg tatgcccggc ccgagttgtt cttggccggc 120
 gtcaataacgg gataataccg cggccacatag cagaacttta aaagtgtca tcattggaaa 180
 acgttcttgc gggcgaaaaac tctcaaggat cttaccgtgt ttggatccca gttcgtatgt 240
 accccactcgtt gaccccaact gatcttcaggc atctttact ttcaccagcg tttctgggtt 300

agcaaaaaaca ggaaggcAAA atgcccggAAA aaagggAAATA agggcgacAC gaaatgttG 360
aatactcata ctcttccttt ttcaatatta ttGAAGCATT tatcagggtt attgtctcat 420
gagcgatAC atatttGAAT gtatttagAA aaataAAACAA atagggttC cgccacatt 480
tccccggAAA gtGCCACCTG acgtcgacGG atcgggAGAT cttccgatCC cctatgttgc 540
actctcagTA caatctgCTC tgatGCCGA tagttaAGCC agtatctgt ccctgtttgt 600
gtgttggagg tcgctgagTA gtgcgcgAGC aaaatttAAAG ctacaacaAG gcaaggcttG 660
accgacaatt gcatgaagAA tctgcttagG tttaggcgtt ttgcgtgtC tcgcgtatgtA 720
cggggccagat atacgcgttG acattgatta ttgactagtT attaatAGTA atcaattacG 780
gggtcattAG ttcatagccc atatatggAG ttccgcgttA cataacttAC ggttaatggc 840
ccgcctggct gaccgcccAA cgacccccGC ccattgacgt caataatgac gtatgttccc 900
atagtaacgc caataggGAc ttccattGA cgtcaatggG tggagtattt acggtaaact 960
gcccacttgg cagtacatCA agtgtatcat atgccaAGTA cgccccctat tgacgtcaat 1020
gacggtaaAT ggcccgcctG gcattatGCC cagtacatGA ctttatggGA ctttcctact 1080
tggcagtgaca tctacgtatt agtcatcgCt attaccatgg tgatgcgtt ttggcagtac 1140
atcaatgggc gtggatagcG gtttgactCA cggggatttC caagtctcca cccattgac 1200
gtcaatggga gtttggggA gCACCAAAAT caacgggACT ttccaaaAtT tgtaacaac 1260
tccggcccat tgacgcaatA gggcggtAG cgtgtacgtt gggaggtcta taataggcaga 1320
gtctctggc taactagaga acccactgCT tactggcttA tcgaaattAA tacgactcac 1380
tatagggaga cccaagctgg cttagcgttA aacttaaAgCT tgccaccatG aagctactgt 1440
cttctatcga acaagcatgc gatatttgc gactttAAAA gctcaagtgc tccaaagaaa 1500
aacogaagtG cgccaagtGT ctgaagaACA actgggagtg tcgctactt cccaaacaca 1560
aaaggcttcc gctgactagg gcacatctGA cagaaggTGA atcaaggcta gaaagactgg 1620
aacagctatt tctactgttA tttcctcgAG aagaccttGA catgattttG aaaatggatt 1680
ctttacagga tataaaAGCA ttgttaacAG gattattttG acaagataat gtgaataaaag 1740
atgcgtcac agatagattG gttcagtgG agactgatAT gcctctaaca ttgagacagc 1800
atagaataag tgcgacatCA tcatcgAAAG agagtagtAA caaaggtaa agacagttGA 1860
ctgtatcgcc gggtaacctt gtgggtcttA aaaaatgAGC atgtgggtcc tcctatggtc 1920
caggttaata ccagactttG ataaaggttCt atggtagaAC atgtttccct ggcacctgtc 1980
ctccctgcct gacagactCA agctttctt ttgttcttta ttttattttt ttatTTTTA 2040
agatggattc tcattctgtc gcccaggctG gagttcagtG acatgatatt ggctcaCTGc 2100
atcccaggTT tcaagcgatt ctccctgcctC aggtcacAA attttgtatt tttagtagAG acagggtttc atgatggtg 2220
ctgccaggac gccaactAA caaggcttgtt ctcaaatttC tgacctaAG tgatctgttC gcctcagcct cccaaagtc 2280
gggattacag gcatgggcca ctgcaccagg ccttcttttG tttttatCA cttcccccatt 2340
ctccaggggc atcatgttA ttttctgcct gtggagttC actggcaag ctcccttca 2400
aagaatcagt cagtgagcT ttccccAACT ttacagctG ctgtctgacc acaggtaatC 2460
taagaaactg cttagattA atgaactggg acaatgtAA ggcaatttct aagtgtatgtt 2520
tctgttagatt aagagggcC ttattaaATA aataccatGA aacaaaaAA aatgactaa 2580
ataggaattc cagggacaat ttaaagggtG ctgggggcca ggtgattatt ttttagacag 2640
aagagtcatT ccaggTctcT tacaacttAC ttgcttaggCc ttgccttact tggctcacac 2700
ttagattaac cctcgTTatG agccgcttaat aatgtgtcA acttgtgatt tcatgaacag 2760
aatgttattt accttattaa tattttggA tggaaattttC catttgatC aacataagaa 2820
gtgtatgtca ttttttttA ttttaatttC tttttttttt ttgagacaga gtctcaCT 2880
gttgcTcagg ctggagtata gttgtggat ctccggtcAc tgccagctt cgcggccagg 2940
ttcaagcgat tctctgtcct cagccttctG agtagctggG attacaggca cctgcccacca 3000
agcaggactA attttggtG ttttagtagA aacgggggttG caccatTTG gctaggctgg 3060
tcttgaactC ctgacctgtG gatccatctG cctcagtaAG ccacctgcct tggcctccca 3120
aagtgcTggG attacaggcG tgagccactG cggccagccG tatctgtcat ttaacaata 3180
aaatccttaA ccctgcattC tcggaaAGAA tgagcacatt aaccatattt ttcctataca 3240
cttgatcaat ttgttatcaa tttaaaaaAT ctgatatcAG gatttcataG aacgcaaaACT 3300
cctcctgcAt ggttggagat ggcgaagtCA gttgcccAGA cttgttaagt tcaggatgca 3360
gcctgtgaaa ggagaggcAA tcctacctGG aagtaccttG gcccacttCA cgacttggat 3420
catctgtttG cctgctaAGC ggttggcgt ggagagcAGA ttttcggctG tatctggtt 3480
tgagctgtca tagcctgtcat atacaatttC agttcaatG tttcaagGA ccataacggg 3540
ggaagggtgtG agcgctcgTG agattgtggA gagctgagGA accagtgcTG tggtgaccGA 3600
gggttctttG gcaggagcGA tgcgttGt cccttccttC gggcttgcG ggggtgggg 3660
tgggggtggG ggctgtgtC gctgtggctG ctccctgtGA atcccttct tttcatctt 3720
tgcattactt ttcttgcTTT catcaccAAAC ctatttctCA tctgcctcCT cccctactGA 3780

aaccctgtt ttccatttgc cagagagac catctgtcag aatatggctg tagttactat 3840
 cataattcaa atattccctt ttaagtattt ctgacccaa caattacaaa gaagcactgc 3900
 tagataaaaa cataaataaa atgtgcta at gaggatgca aacagtatta aaatgtaaac 3960
 taaaatagg aaggaatttg tggccaaaga aaaactgatt acgattctt cagtccttta 4020
 aaatgcatca taaaactaaa aatgtttgc tacagaataa cagttggct tcttttatga 4080
 tggcaaaaga ggttaattac gattgtcatg aaaacacatt gagcagctgt atactgttaa 4140
 caaaaacc attatatgag ctcgtaaaag ctggcactt atatattcac atttcctaga 4200
 aattagtt tcttccaaat tggagtcgt accaaaagag actgtttct attaattgtg 4260
 gcttttacc agtcaattt cagagccaac tggatcatct ttcttttggg atatatccca 4320
 catattat tggaaattt atataactg attatataac actcagaagc atacagtata 4380
 atgttgcataa gagtaatttt ttaagttccc agagatctgt aaatttttaa cgttaaaaaaa 4440
 tgccaggttc agtagattct ttcaccaacg tacatctgtt tagaaattac tgaaagaaaat 4500
 cataacgcataactctgcag tttattttat gaaggctaat taataggtac ttactcatta 4560
 aagactaggt ctggcataa atagagaaat tgctgttcg tatgtttgtt cgatctccag 4620
 ctcaaggcaa atgatgatag acacatccaa gaataactgga tttaggtaat ttggcctca 4680
 agaggcaagt tttaaatcc tgaagaacaa aacaattaat cacagaaata cacttagcat 4740
 ttaagtacat tccaggaaga tgcttcttaa gtcaacccccca aacagccacc ttctttcac 4800
 tttctctt tctgacagac tcaaataat tcaacagtca ttacagagt gcccaccact 4860
 ggcccagtcc tggccacat ttcaattttt attcatccta aattcttagat gagggtaaac 4920
 tgccaggggca aactaaaattt gcaaagttca gcaagttca gtgtggatga cagcctaaaaa 4980
 cccttggta gattgagaga atttgttata acctgtaaaaa cttgcttcgt cccaccacca 5040
 attaataaca cttaaaataaa ggttaatttataactgtca aaccctgtgt atctgcaata 5100
 atgggtgggag atttcattgt tcatgaggcg gccgctcgag tctagaggc cctgtttaaac 5160
 ccgctgatca gcctcgactg tgcctcttag ttgccagcca tctgttgtt gcccctcccc 5220
 cgtgccttcc ttgaccctgg aaggtgccac tcccactgtc ttctcttaat aaaatgagga 5280
 aattgcattcg cattgtctga gttaggtgtca ttctattctg ggggggtgggg tggggcagga 5340
 cagcaagggg gaggattggg aagacaatag caggcatgtc gggatgcgg tgggctctat 5400
 ggcttctgag gggaaagaa ccactgtggg ctctaggggg tatccccacg cccctgttag 5460
 cggcgatca agcgccggg gtgtgggtt tacgcgcage gtgaccgceta cacttgcac 5520
 cggccctagcg cccgctcctt tcgttctt cccttcctt ctcgcgcacgt tcgcccgtt 5580
 tccccgtcaa gctctaaatc gggggctccc tttaggggtt cgatttagt ctttacggca 5640
 ctcgcacccc aaaaaacttg attaggggttga tgttcacgt agtggccat cggccctgata 5700
 gacggttttt cggcccttga cgttggagtc cacgttctt aatagtggac tcttgttcca 5760
 aactggaaaca acactcaacc ctatctcggt ctattcttt gatttataag ggattttgcc 5820
 gatttcggcc tattggtaa aaaaatgagct gatttacaa aaatttaacg cgaattaatt 5880
 ctgtggatg tggcgttgcgtt agggtgttga aagtccccag gctcccccacg aggccagaagt 5940
 atgcaaaagca tgcattctcaa ttatgtcagca accaggtgtt gaaagtcccc aggtccccca 6000
 gcaggcagaa gtatgcaaaag catgcatttc aatttgcgtt caaccatagt cccgcccccta 6060
 actccggccca tcccgccccctt aactccgccc agttccgccc attctccgccc ccatggctga 6120
 ctaattttttt ttatattgc agaggccgag gccgcctctg cctctgagct attccagaag 6180
 tagtgaggag gcttttttgg aggcttaggc ttttgcacaaa agtccccggg agcttgcata 6240
 tccattttcg gatctgatca gcacgtgtt aaaaaggctt aactcaccgc gacgtctgtc 6300
 gagaagttt tggatcgaaaa gttcgacacg gtcgtccgacc tgatgcgtt ctcggaggggc 6360
 gaagaatctc gtgttttgcgtt cttcgatgtt gggggcggtt gatatgttctt ggggttaat 6420
 agctgcgcgg atggtttcta caaagatctgt tatgtttatc ggcacttgc atcgccgcgc 6480
 ctcccgattt cggaaagtgtt tgacatttggg gaattcagcg agagctgtac ctattgcattc 6540
 tcccgccgtt cacagggtgtt cacgttgcac gacctgcctt aaaccgaact gcccgtgtt 6600
 ctgcaggccgg tgcgggaggc catggatgcgtt atcgctgcgg ccgatcttag ccagacgc 6660
 gggttcgcccat tccggacc gcaaggaaatc ggtcaataaca ctatcatggcg tgatttcata 6720
 tgcgcgtt ctgatccccca tggatgttccatc tgccaaactgt tgatggacga caccgtcgt 6780
 gctccgtcg cgaggctt cgtgatgttccatc tgccaaactgt tgatggacga caccgtcgt 6840
 cggcacccctcg tgacgcggat tttcggttccatc aacaatgttccatc tgacggacaa tggccgcata 6900
 acaggcggtca ttgactggatc cgaggcgatc ttcggggattt cccaaatatacg ggtcgccaaac 6960
 atcttcttctt ggaggccgtt gttggcttgcgtt atggagcgcg agacgcgcata cttcgagcgg 7020
 aggcatccgg agcttgcagg atcgccgcgg ctccggccgtt atatgtcccg cattggctt 7080
 gaccaactctt atcagagctt ggttgcaggc aatttcgtatc atgcagctt ggcgcagggt 7140
 cgatgcgcgcg caatcgatccg atccggagcc gggactgtcg ggcgtacaca aatcgcccg 7200
 agaaggcgccg cggctctggac cgtggctgtt gtagaagttac tcgcccgtt gggaaaccga 7260

cgccccagca ctcgtccgag ggcaaaggaa tagcacgtgc tacgagattt cgattccacc 7320
 gcccgccttct atgaaaagggtt gggcttcgga atcgtttcc gggacgcggg ctggatgatc 7380
 ctccagcgcg gggatctcat gctggagttc ttccgcacc ccaacttgtt tattgcagct 7440
 tataatggtt acaaataaag caatagcatc acaaattca caaataaagc attttttca 7500
 ctgcattcta gttgtggttt gtccaaactc atcaatgtat cttatcatgt ctgtataccg 7560
 tcgacctcta gtagagctt ggcgtaatca tggcatagc tggttccgtgt gtgaaattgt 7620
 tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa agcctgggt 7680
 gccttaatgg tgagctact cacattaat gcttgcgcg cactgcccgc ttccagtcg 7740
 gggaaacctgt ctgcgcagct gcatattaatga atccggccaaac ggcggggag aggccgtttg 7800
 cgtattggc gctttccgc ttccctcgctc actgactcgc tgcgcgtgt cgttcggctg 7860
 cggcgagcgg tattcagctca ctcaaaggcg gtaatacggg tattccacaga atcagggat 7920
 aacgcaggaa agaacatgtg agaaaaaggc cagcaaaaagg ccaggaaccg taaaaaggcc 7980
 gcgttgcgg cgttttcca taggctccgc cccctgacg agcatcaca aaatcgacgc 8040
 tcaagtctaga ggtggcgaaa cccgacagga ctataaaagat accaggcggtt tccccctgga 8100
 agctccctcg tgcgctctcc tggccgacc ctgcccgtta coggataacct gtccgcctt 8160
 ctccctcgg gaagcgtggc gcttctcat agctcacgct gtaggtatct cagtcgggt 8220
 taggtcggtc gctccaagct gggctgtgtg caagaaaccc cgttccagcc cgaccgctgc 8280
 gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg 8340
 gcagcagcca ctggtaacag gattagcaga gccgaggatg taggcgggtgc tacagagttc 8400
 ttgaagtggt ggcttaacta cggctacact agaagaacag tatttggat ctgcgctctg 8460
 ctgaaggccag ttaccttcgg aaaaagagtt ggtagctt gatccggcaa acaaaccacc 8520
 gctggtagcg gttttttgtt ttgcaagcagc cagattacgc gcagaaaaaa aggatctcaa 8580
 gaagatccctt tgcgttttc tacggggctt gacgctcagt ggaacgaaaa ctcacgttaa 8640
 gggattttgg tcatgagatt atcaaaaagg atcttcaccc agatcccttt aaattaaaaa 8700
 tgaagttta aatcaatcta aagtatatat gagtaaactt ggtctgacag ttaccaatgc 8760
 ttaatcagtg aggccacccat ctcagcgtc tgcgttatttc gttcatccat agttgcctga 8820
 ctccccgtcg ttagataaac tacgatacgg gagggtttac catctggccc cagtgcgtca 8880
 atgataccgc gagacccacg ctcaccggct ccagatttat cagcaataaa ccagccagcc 8940
 ggaaggcccg aegcgcagaag tggccctgca actttatccg cctccatcca gtctattaaat 9000
 tggccggg aagcttaggt aagtagttcg ccagttataa gtttgcgcaa cgttgttgc 9060
 attgctacag gcatcgtgtt gtcacgcgtc tggccgtt cagtcgggt 9120
 tcccaacgat caaggcgagt tacatgatcc cccatgtt gcaaaaaaagg ggttagctcc 9180
 ttccggcctc cgatcgtgtt cagaagtaag tt 9212

<210> 79
 <211> 10517
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 pKI-Gal-DBD-MR sequence

<220>
 <221> modified_base
 <222> (9953)..(9953)
 <223> any nucleotide

<400> 79
 atcggacgat tgcgtcgcat cgaccctcgcc cccaagctgc atcatcgaaa ttgccgtcaa 60
 ccaagctcg atagagttgg tcaagaccaa tgcggagcat atacgcccgg agccgcggcg 120
 atcctgcaag ctccggatgc ctccgcgtca agtagcgcgt ctgcgtctcc atacaagccaa 180
 accacggcct ccagaagaag atgttggcga cctcgtattt ggaatccccg aacatcgcc 240
 cgctccagtc aatgaccgct gttatgcggc cattgtccgt caggacattt tggagccga 300
 aatccgcgtg cacgaggtgc cggacttcgg ggcagtccctc ggcccaaaggc atcagctcat 360
 cgagagcctg cgccgacggac gcactgacgg tgcgtccat cacagttgc cagtgataca 420
 catggggatc agcaatcgcg catatgaaat caccatgtt agtgtattga ccgattccct 480

gcggccggaa tggggccgaac ccgctcgat gcgtaagat ggcgcagcg atcgcattca 540
tggcccccgc gaccggctgc agaacaggcg gcagttcggt ttcaggcagg tcttgcacacg 600
tgacacctg tgacaggcg gagatcaat aggtcaggct ctcgctgaat tccccatgt 660
caagcacttc cggaatcggt agcgcggccg atgcaaagtgc ccgataaaca taacgatctt 720
tgtagaaacc atcggcgacg ctatcaccc gcaggacata tccacccct cctacatcga 780
agctgaaagc acgagattct tgcgcctcg agagctgcac caggtcggag acgctgtcga 840
actttcgat cagaaacttc tcgacagacg tcgcggtag ttcaggttt ttcatgcagg 900
tcgaaggccc ggagatgagg aagaggagaa cagcgcggca gacgtcgct tttgaagcgt 960
gcagaatgcc gggtccggag gaccttcgcg cccgcggccgc ccctgagccc gcccctgagc 1020
ccgccccccg acccaccctc tcccagccctc tgagcccaaga aagcgaagga gccaagctgc 1080
tattggccgc tgccccaaag gcctaccgc tccattgtct cagcgtgtct tgcacatctg 1140
cacgagacta gtgagacgtg ctacttccat ttgtcagtc ctcacgacg cgagctgcgg 1200
ggcgaaaaaaa aacttcctga ctaggggagg agtagaaggt ggcgcgaagg gccaccaaaa 1260
gaagggagcc ggttggcgct accgggtggat gtggaatgtg tgcgaggcga gaggccactt 1320
gtgatgcgcc aagtgcgcg ggggctgtca aagcgcatgc tccagactgc cttggaaaaa 1380
gcgcctcccc taccggtag aattccgtca gcccgggggtt gacattgatt attgactagt 1440
tattaatagt aatcaattac ggggtcatta gtcatagcc catatatggc gttccgcgtt 1500
acataactt cgttaaatgg cccgcctggc tgaccggccca acgaccccg cccattgacg 1560
tcaataatga cgtatgttcc catagtaacg ccaataggga ctttcattt acgtcaatgg 1620
gtggagtt tacggtaaac tgccacttg gcagatcatc aagtgtatca tatgccaatg 1680
acgcccccta ttgacgtcaa tgacggtaaa tgccccggctt ggcattatgc ccaagtatcg 1740
accttatggg actttccatc ttggcagttt acatcacgtat tagtcatcgc tattaccatg 1800
gtgatgcggc tttggcagta catcaatggg cgtggatagc ggtttagtct acgggggattt 1860
ccaagtctcc accccattga cgtcaatggg agttgtttt ggcacccaaa tcaacgggac 1920
tttccaaaaat gtcgtAACAA ctccgcggccca ttgacgcaaa tggcggttag gcgtgtacgg 1980
tgggagggtct atataagcag agctctctgg ctaacttagag aacccactgc ttactggctt 2040
atccggatcc actagttcta gcgtttaaac ttaagcttgc caccatgaag ctactgtctt 2100
ctatcgaaca agcatgcgtat atttgcgcac tttaaaaagct caagtgtcc aaagaaaaac 2160
cgaagtgcgc caagtgtctg aagaacaact gggagtgtcg ctactctccc aaaacccaaa 2220
ggtctccgct gactaggcgca catctgacag aagtggaaatc aaggcttagaa agactggAAC 2280
agctattttt actgattttt cctcgagaaag accttgacat gattttgaaa atggattctt 2340
tacaggatataaaggcatttgcgca tagattggct tcagtggaga ctgatatgccttcaacatttgc 2400
ccgtcacaga tagattggct tcggaaagaga gttagtaacaa aggtcaaaga cagttgactg 2460
gaataagtgc gacatcatca tacctttgtg gtgtttaaaa aatgagcatg tgggcctcc tatggccag 2520
tatcgccggg gacttttgcgat gactttgata aagttctatg gttagaacatg ttccctggc acctgtcc 2580
gttaatacca gactttgata aatccctgtac agactcaago tcttcttttgc ttttttatttatttatttataa 2640
cctgcctgac agactcaago tggatttcata ttctgtcgcc caggctggag ttctgtgaca tgatattggc tcactgcattc 2700
ccaggtttca agcgattctc ccaggacgccc caactaaatt ttgttattttt agtagagaca ggttcatg atggtgccaa 2760
ggctggcttc aaattccgtaa caccaggcct tcttttgc tttatcacct ccccatctc 3000
attacaggca tggggccactg cccaaactttt acagctgtcg tctgaccaca gggcaagctc ctcttcaaa 3060
cagggccatc atgttattttt aacttacttg ctggccttgc tttttttttttagacagaaatgactaaataa 3120
aatcagtcag tggagcttcc aactgggaca atgtaatggc aatttctaaatg tggatgttttctt 3180
gaaactgtct tagattaatg ttagattaaat accatgaaac aaaaaaaaaat gactaaaataa 3240
gttagattaaatggatcccg ggacaatggc aagggtgctg gggggccaggat gatttttttttagacagaaatgactaaataa 3300
aacttccatcc ggtcttttgc aacttacttg ctggccttgc ctttacttgg tctacactt 3360
gatatttttgc tttttttttttagacagaaatgactaaataa 3420
gtatatttttgc tttttttttttagacagaaatgactaaataa 3480
tatgtcattt tttttttttttagacagaaatgactaaataa 3540
gctcaggctg gagatatagtg gtggatctc ggctcactgc agcctctgcc gcccagggttc 3600
aagcgatctt cctgcctcgc cttcctgtat agctgggattt acaggcacct gccaccaagc 3660
aggactaattt tttgtgtttttagacagaaatgactaaataa 3720
tgaactccgtt acctcgatgc ccatctgcct cagtaagccaa cctgccttgg cttcccaaa 3780
tgctggattt acaggcgatgc gcccactgc ccaagccgtat ctgtcattttt aacaataaaa 3840
tccttaaccc tgcattctcg gaagaaatga gcacattaaac catatttttgc tatacactt 3900
gatcaatttgc ttagatcaattt aaaaaaaaaatctg atatcaggat ttcatagaac gcaaactctc 3960

cctgcatgg tggagatggc gaagtca gcccagactt gttaagttca ggatgcagcc 4020
 tgtgaaagga gaggcaatcc taccttggaa taccttgc cacttcacga cttggatcat 4080
 ctgtttgcct gctaagcggt tgagcgtgga gaggcatt tcggctgtat ctggtttga 4140
 gctgtcatag cctgcatata caatttcagg ttcaatgttt tcaaggacca taacggggga 4200
 aggtgtgagc gctcggtgaga ttgtggagag ctggaggaacc agtgctgtgt tgaccgaggg 4260
 ttctttgca ggagcgatgt acgttgc ttcctctggg ctttgcgggg gtgggggtgg 4320
 ggggtggggc tgctgtgtc gtggctgtc ctcgtgaatc cttttttt catctttgc 4380
 attactttc ttgtttcat caccaccta ttcttcatct gccttc ttcctccc ctactgaaac 4440
 ccctgtttc ctatgccag agagagccat ctgtcagaat atggctgttag ttactatcat 4500
 aattcaaaaata ttcccttta agtattttt gacccaacaa ttacaagaa gcaactgctag 4560
 atacaacat aaataaaaatg tgctaatgaa gaatgcaaac agtattaaaa tgtaaactta 4620
 aaatagaaag gaatttggc ccaaagaaaa actgattacg atttttcatt tcttttaaaa 4680
 tgcataa aactaaaaat gtttgc ttttgc ttttgc tttatgatgg 4740
 caaagaaggg taattacgtat tgctatgaaa acacattgag cagctgtata ctgttaacaa 4800
 aagaaccatt atatgagctc gtaaaagctg gcaacttata tattcacatt tcctagaaat 4860
 tagattgtct tcccaattgg agtcgatacc aaaagagact gtttctatt aattgtggct 4920
 ttgtaccagt caatttcag agccaaactgg atcatcttc ttttggata tatccacat 4980
 attattatgt ggaatttata tatactgatt tatatacact cagaagcata cagtataatg 5040
 ttgcaaaagag taattttta agtccacaga gatctgtaaa tttttaacgt taaaaaatgc 5100
 cagtttcagt agattcttc accaacgtac atctgttttag aaattactga aagaaatcat 5160
 aacgcataac tctgcagttt attttatgaa ggctaattaa taggtactta ctcattaaag 5220
 actaggctg gtgcaaaaata gagaaattgg ctgttcgtat gtttgc tctccagctc 5280
 aaggcaaatg atgatagaca catccaagaa tactggatta gggtaatttgc tccctcaaga 5340
 ggcaagttt taaatcctga agaacaacaa aattaatcac agaaatacac ttagcattt 5400
 agtacattcc aggaagatgc ttcttaagtc aaccccaaaac agccacctt ctttcaactt 5460
 tctcttttct gacagactca aatcaattca acagtca ttttgc caccactggc 5520
 ccagtcctgt tccacatttc actttaaatt catcttaat tcttagatgag gttaaaactgc 5580
 cagggcaaac taaaatgca aagttcagca agtttcgtg tggatgacag cttaaaaccc 5640
 ttggtagat tgagagaatttgc ttttttttttgc ttttttttttgc ttttttttttgc 5700
 aataacacta aaataaaaggg taatttataa actgtcaaac cctgtgtatc tgcaataatg 5760
 gtgggagatt tcatttttca tgaggccggc gcttgc ttttttttttgc ttttttttttgc 5820
 tactgtccgc aaggggccggc gctatttttgc ttttttttttgc ttttttttttgc 5880
 cgccaaactc ttcatcctt ttttttttttgc ttttttttttgc ttttttttttgc 5940
 tctctttttt tttttttttt ttttttttttgc ttttttttttgc ttttttttttgc 6000
 aaatgctgca ctgtgcggca aggccgtga ataaaacgact cgaatccat cccccccgc 6060
 cgttctaaa gttgcctttt ttttttttttgc ttttttttttgc ttttttttttgc 6120
 ggcgcaaccc ccagccactg ttttttttttgc ttttttttttgc ttttttttttgc 6180
 agctgttcac cgggggtggc ttttttttttgc ttttttttttgc ttttttttttgc 6240
 agttcagcgt gtccggcgag gtttttttttgc ttttttttttgc ttttttttttgc 6300
 tcatttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 6360
 acggccctgca gtgtttcgcc ttttttttttgc ttttttttttgc ttttttttttgc 6420
 cccatcgcc cgaaggctac ttttttttttgc ttttttttttgc ttttttttttgc 6480
 acaagaccccg cggccgggtg aagttcgagg gtttttttttgc ttttttttttgc 6540
 agggcatcga ttcaaggag gtttttttttgc ttttttttttgc ttttttttttgc 6600
 acagccacaa cgtctatatac ttttttttttgc ttttttttttgc ttttttttttgc 6660
 agatccgcca caacatcgag gtttttttttgc ttttttttttgc ttttttttttgc 6720
 cccatcgcc cggccggccgc ttttttttttgc ttttttttttgc ttttttttttgc 6780
 ccctgagcaa agaccccaac gtttttttttgc ttttttttttgc ttttttttttgc 6840
 cccatcgcc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 6900
 atggttacaa ataaagcaat gtttttttttgc ttttttttttgc ttttttttttgc 6960
 attcttagtt tggtttgtcc ttttttttttgc ttttttttttgc ttttttttttgc 7020
 cccggccgca cggccgggtggc ttttttttttgc ttttttttttgc ttttttttttgc 7080
 actggccgtc gtttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7140
 ctttgcagca catccccctt ttttttttttgc ttttttttttgc ttttttttttgc 7200
 cccttccaa cagttgcgca gtttttttttgc ttttttttttgc ttttttttttgc 7260
 aagcgccggcg ggtgtgggtgg ttttttttttgc ttttttttttgc ttttttttttgc 7320
 gcccccttctt ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7380
 agctctaaat cgggggttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7440

caaaaaaactt gattagggtg atggttcacg tagtgggccca tcgcccgtat agacggttt 7500
 tcgccttgc acgttgagt ccacgtctt taatagtggc ctctgttcc aaactggAAC 7560
 aacactcaac cctatctcg tctattctt tgatttataa gggatttgc cgatttcggc 7620
 ctattggtta aaaaatgagc tgattaaca aaaatttaac gcgaattta acaaataatt 7680
 aacgcttaca atttaggtgg cactttcgg gaaaatgtgc gcggAACCC 7740
 ttttctaaa tacattcaa tatgtatccg ctcatgagac aataaccctg ataaatgctt 7800
 caataatatt gaaaaggaa gatgtatgat attcaacatt tccgtgtcg ctttattccc 7860
 tttttcgcc cattttgcct tectgtttt gtcacccag aaacgcttgc gaaagtaaaa 7920
 gatgctgaag atcagtggg tgcacgagtg gtttacatcg aactggatct caacagcgg 7980
 aagatccttgc agagtttcg ccccgaagaa cgttttccaa tggatgagcac ttttaaagtt 8040
 ctgctatgtg ggcgggtatt atccccgtt gacggccggc aagagacaact cggtcgccc 8100
 atacactatt ctcagaatga cttgggttag tactcaccag tcacagaaaa gcatcttacg 8160
 gatggcatga cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg 8220
 gccaacttac ttctgacaac gatcgagga cccaaggagc taaccgcTTT tttgcacaac 8280
 atggggatc atgtaactcg ctttgcgtt tggaaccgg agctgaatga agccatacca 8340
 aacgacgagc gtgacaccac gatgcgttgc gcaatggca caacgttgcg caaactatta 8400
 actggcgaac tacttactct agcttcccg caacaattaa tagactggat ggaggcggat 8460
 aaagttgcag gaccacttct ggcgtcgcc cttccggctg gctggTTT tgctgataaa 8520
 tctggagccg gtgagcgtgg gtctcgccgt atcattgcag cactggggcc agatggtaag 8580
 ccctcccgta tctgttttat ctacacgacg gggagtcagg caactatggta tgaacgaaaat 8640
 agacagatcg ctgagatagg tgcctcactg attaaggattt ggttaactgtc agaccaagtt 8700
 tactcatata tacttttagat tgatttaaaa cttcattttt aatttaaaag gatcttaggt 8760
 aagatcctt ttgataatct catgacccaa atcccttaac gtgaggTTT gttccactga 8820
 ggcgtcagacc ccgtagaaaa gatcaaaaggatc tcttcttgcg atccttttt tctgcgcgt 8880
 atctgcgtct tgcaaaacaaa aaaaccaccc ctaccagcgg tggTTTgTTT gccggatcaa 8940
 gagctaccaa ctcttttcc gaaggtaact ggcttcagca gagcgcagat accaaataact 9000
 gtccttcttag ttagccgtt gttaggccac cacttcaaga actctgttagc accgcctaca 9060
 tacctcgctc tgcataatctt gttaccagt gctgcgtcgc gttggcgtt gtcgtgtctt 9120
 accgggttgg actcaagacg atagttaccg gataaggcgc aacgggtcggg ctgaacgggg 9180
 ggttcgtcga cacagcccag cttggagcga acgacctaca cggaaactgag atacctacag 9240
 cgtgagctat gagaagcgc cacgcttccc gaagggagaa aggccggacag gtatccggta 9300
 agcggcaggg tcggaacacgg agagcgcacg agggagctt cagggggaaa cgcctggat 9360
 ctttatagtc ctgtcggtt tcgcccaccc tgcattttt gtcgattttt gtgtatgtcg 9420
 tcaggggggc ggagcctatg gaaaaacgcg acgaacgcgg ctttttacg gttccctggcc 9480
 ttttgcgtgc cttttgcata catgtttttt cctgcgttac cccctgattt tggatgat 9540
 cgtattaccg cttttgagtg agctgatacc gtcgcgcgc gccgaacgc cgagcgcagc 9600
 gagtcagtga gcgaggaagc ggaagagcgc ccaatacgcg aaccgcctt cccgcgcgt 9660
 tggccgattc attaatgcag ctggcacgcg aggtttcccg actggaaagc gggcagttag 9720
 cgcaacgcaaa ttaatgttag ttagctact cattaggcgc cccaggctt acactttatg 9780
 cttccggctc gtatgttgtg tggaatttgtg agccgataac aatttcacac agggaaacagc 9840
 tatgaccatg attacgcac ggcgcacatt aaccctcact aaagggaaaca aaagctgggt 9900
 accggggcccc ccctcgaggt cgacggatc gataaggctt cagtggatca ganaaacaat 9960
 gatcccacga gagatctata gatctataga tcatgagtgg aggaatgagc tggcccttaa 10020
 tttggTTTtgc tttttttttttaaa ttatgatatac caactatgaa acattatcat aaagcaatag 10080
 taaagagcct tcaatggaa gcaaggattt atctaattcc acccccaccc ccaccccccgt 10140
 agctccaatc cttccattca aaatgttagt actctgttctt cacccttctt aacaaagtat 10200
 gacagtaaaa aacttccatt ttagtggaca tctttattgt ttaatagatc atcaattttat 10260
 gcagacttac agctgggtt tcatcagatt gccatgtcg gtcagaaca gtgagtgacg 10320
 aagctaacca gaggctacat acgtcagaga gagagcttag cctttacagc tcacttcctt 10380
 ttcaggcat tctttggcc tcggacgat gctggggcgat cgggtttccac ttcgtgcgag 10440
 tacttctaca cagccatcg gtcagacggc cgcgccttcg cgggcgattt gtgtacgccc 10500
 gacagttcccg gtcggc 10517

```

<210> 80
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

<400> 80
gatctagaga gccacacccc atgttgtctc ac          32

<210> 81
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

<400> 81
gagcggccgc ttgccaaggg ccctgtgaag cagg          34

<210> 82
<211> 5351
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      pKI-CMV-SD sequence

<400> 82
tcccttagt gagggttaat tgcgcgcttg gcgtaatcat ggtcatagct gtttcctgtg 60
tgaattgtt atccgctcac aattccacac aacatacggag ccgaaagcat aaagtgtaaa 120
gcctgggtg cctaattgagt gagctaactc acattaattt cgttgcgtc actgcccgt 180
ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg cgccccggaga 240
ggcggtttgc gtattggcg ctcttcggct tcctcgctca ctgactcgct gcgctcggtc 300
gttcggctgc ggcgagcggt atcagctcac tcaaaggcggt taatacggtt atccacagaa 360
tcagggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt 420
aaaaaggccg cgttgttgc gttttccat aggtccgccc cccctgacga gcatcacaaa 480
aatcgacgct caagtccagag gtggcgaaac ccgcacaggac tataaagata ccaggcggtt 540
ccccctggaa gtccttcgt ggcgttcgt gttccgaccg tgccgttac cggataacctg 600
tccgccttc tccttcggg aagcggtggcg ctttcgtcata gctcacgctg taggtatctc 660
agttcggtgt aggtcggtcg ctccaaagctg ggctgtgtgc acgaaacccc cggtcagccc 720
gaccgctcgcc ctttatccgg taactatcgat cttgagtcac accccggtaa acacgactta 780
tcggcactgg cagcagccac tggtaacagg attagcaggag cgaggtatgt aggccgtgt 840
acagaggatctc tgaagtgggtg gcctaactac ggctacacta gaaggacagt atttggatc 900
tgcgtctgc tgaagccagt taccttcggg aaaagagttt gtagctttt atccggcaaa 960
caaaccaccc ctggtagccgg tggttttt gtttgcggc acgagattac ggcggaaaa 1020
aaaggatctc aagaagatcc tttgatcttt tctacgggggt ctgacgctca gtggAACGAA 1080
aactcactgtt aagggtttt ggtcatgaga tttcaaaaaa ggatcttcac ctatgttcatt 1140
ttaaataaaa aatgtttt taaatcaatc taaatgtat atgatgtttt ttgggtctgac 1200
agttaccaat gcttaatcgt tgaggcacct atctcagcgtc tctgtctatt tcgttcatcc 1260
atagttgcct gactccccgt cgtgttagata actacgatac gggagggtt accatctggc 1320

```


tgcggcggag	gcatccggag	cttgcaggat	cggcgccggt	ccggggcgat	atgctccgca	4860
tttgtcttga	ccaactctat	cagagcttgg	tttgacggcaa	tttcgatgtat	gcagcttggg	4920
cgcagggtcg	atgcgacgca	atcgctcgat	ccggagccgg	gactgtcggg	cgtacacaaa	4980
tcgcccgcag	aagcgccgco	gtctggaccg	atggctgtgt	agaagatactc	gccgatagtg	5040
gaaaccgacg	ccccagcaact	cgtccgaggg	caaaggaatg	agtcgctgat	tcagcctcgat	5100
ctgtgccttc	tagttgccag	ccatctgttg	tttgcccttc	ccccgtgcct	tcccattgtc	5160
tgagtaggtg	tcattctatt	ctgggggttg	gggtggggca	ggacagcaag	ggggaggatt	5220
ggaaagacaa	tagcaggcat	gctggggatg	cggtgggctc	tatggcttct	gaggcggaaa	5280
gaaccagctg	ggggatatac	agcttatcga	taccgtcgac	ctcgaggggg	ggcccggtac	5340
ccagcttttg	t					5351

```
<210> 83
<211> 9131
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      pKI-CMV-SD-Vanilloid sequence
```

<400> 83
cgatcgtctgc ggccgatctt agccagacga gcgggttcgg cccattcga ccgcaaggaa 60
tcggtcaata cactacatgg cgtgattca tatgcgcgtat tgctgatccc catgtgtatc 120
actggcaaac tgtgtatggac gacaccgtca gtgcgtccgt cgccgaggct ctcgatgagc 180
tgtatgcttg ggccgaggac tgccccaaag tccggcacct cgtgcacgcg gatttcggct 240
ccaacaatgt cctgacggac aatggccgca taacagcggt cattgactgg agcgaggcga 300
tgttcgggaa ttcccaatac gaggtcgcca acatcttctt ctggaggccg tggttggct 360
gtatggagca gcaagacgcg tacttcgagc ggaggcatcc ggagcttgcg gatgcgcgc 420
ggctccgggc gtatatgctc cgcattggtc ttgaccaact ctatcagagc ttgggtgacg 480
gcaatttcga tgtatgcact ggccgactgt cggcgatcaca caaatcgccc gcaagaagcgc ggccgtctgg accgatggct 540
tgttagaaagt actcgccat agtgaaacc gacgccccagg cactcgatccg agggcaaagg 600
aatgagctcg ctgtatcagcc tcgactgtgc cttctagttt ccagccatct gttgtttgcc 720
cctcccccgt gccttcccat tgtctgagta ggtgtcattt tattctgggg ggtgggttgg 780
ggcaggacag caagggggag gattggaaag acaatagcag gcatgtggg gatgcgttgg 840
gctctatggc ttctgaggcg gaaagaacca gctgggggat atcaagctt tcgataccgt 900
cgacctcgag gggggggcccg gtaccagct ttgttccctt ttagtgaggg ttaattgcgc 960
gcttggcgta atcatggtca tagctgttcc tctgtgtgaaa ttgttatccg ctcacaattc 1020
cacacaacat acgagccgga agcataaaagt gtaaagcctg gggtgctaa tgagtgagct 1080
aactcacatt aattgcgttgc cgctcaactgc ccgccttcca gtcggaaac ctgtcgtgcc 1140
agctgcattt atgaatcgcc caacgcgcgg ggagaggcgg tttgcgtatt gggcgctt 1200
ccgcttcctc gctcaactgac tcgctgcgtcg gctgoggcga gcggtatcag 1260
ctcaactcaaa ggccgtaata cggttatcca cagaatcagg ggataacgcg gaaagaaca 1320
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaaa ggccgcgttgc ctggcgtttt 1380
tccatagggct ccggcccccgt gacgagcatc aaaaaatcg acgctcaagt cagaggtggc 1440
gaaaccggac aggactataa agataccagg cgtttcccccc tggaaagctcc ctctgtcgct 1500
ctccctgttcc gaccctgccc cttaccggat acctgtccgc ctttctccct tcgggaaagcg 1560
tggcgcttc tcatagctca cgctgttaggt atctcagttc ggtgttaggtc gttcgcctca 1620
agctgggctg tgtgcaccaa ccccccgttcc agccgcacccg ctgcgcctta tccggtaact 1680
atcgcttgcgttca gtcacaaaaac gtaagacacg acttatecgcc actggcagca gcaactggta 1740
acaggatttag cagagcgagg tatgttaggcg gtgtacaga gttctttaag tgggtggctt 1800
actacggcta cactagaagg acagtatttg gatatctgcgc tctgctgaag ccagttacct 1860
tcggaaaaaaag agttggtagc tcttgatccg gcaaaacaaac caccgcgttgc agcgggtgg 1920
tttttggtttcaagcagcagcattacgcga gaaaaaaaaagg atctcaagaa gatcccttgc 1980
tcttttctac ggggtctgac gctcagtggaa acgaaaaactc acgcttaagggtt attttggct 2040
ttagattatc aaaaaggatc ttcacccataga tcccttttaaa taaaaatga agttttaaat 2100
caatctaaag tatatatgat taaacttggt ctgacagttt ccaatgcctt atcagtggagg 2160

cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc cccgtcgtgt 2220
 agataactac gatacgggag ggcttaccat ctggccccag tgctgaatg ataccgcgag 2280
 acccacgctc accggctcca gatttatcg caataaaacca gccagccgga agggccgagc 2340
 gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt tgccggaaag 2400
 ctagagaag tagttcgcca gttaatagtt tgogcaacgt tggtgccatt gctacaggca 2460
 tcgtgggtgc acgctcgctg ttggtatgg cttcattcgat ctccggttcc caacgatcaa 2520
 ggcgaggtaatgatcccccc atgttgtaaaaagcggt tagctccctt ggtcctccga 2580
 tcgttgtaaagttg gcccgtgt tatcattcat gtttatggca gcaactgcata 2640
 attctttac tgcatttttgc tccgtaatgt gctttctgt gactggtgag tactcaacca 2700
 agtcatctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg tcaatacggg 2760
 ataataccgc gcacatagc agaactttaa aagtgtcat cattggaaaa cgttttcgg 2820
 ggcgaaaact ctcaggatc ttaccgtgt tgagatccag ttcgatgtaa cccactcg 2880
 caccctactg atcttcagca tcttttactt tcaccagcg ttcgtggtaa gcaaaaacag 2940
 gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgtga atactcatac 3000
 tcttccttt tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggataca 3060
 tatttgaatg tatttagaaaa aataaaacaaa taggggttcc ggcacattt cccgaaaag 3120
 tgccacctaattttaaattttt gttaaaattt gctttaaattt tttgttaat 3180
 cagctcattttttaaccat aggccgaaat cggcaaaatc ctttataat caaaagaata 3240
 gaccgagata gggttgagtg ttgttcagt ttggacaacag agtccactat taaagaacgt 3300
 ggactccaac gtcaaaaggc gaaaaaccgt ctatcaggc gatggccac tacgtgaacc 3360
 atcaccctaa tcaagttttt tggggtcgag gtccgtaaa gcaactaaatc ggaaccctaa 3420
 agggagcccc cgatttagag cttgacgggg aaagccggcg aacgtggcga gaaaggaagg 3480
 gaagaagcg aaaggagcgg ggcgttagggc gctggcaagt gtacggcgtca cgctgcgcgt 3540
 aaccaccaca cccgcccgc ttaatgcgc gctacaggc gcttccattt cgccatttc 3600
 gtcgcgaac tggggaaag ggcgttcggt gccggccctc tcgttattac gccagctggc 3660
 gaaaggggaa tggctgcaaa ggcgtttaag ttggtaaag ccagggtttt cccagtca 3720
 acgtttaaa acgacggcca gtgagcgc gtaatacgcac tcactatagg gcaatttgg 3780
 gtcaccgc ggtggccggc gcttgcgcaag ggcctgtga agcagggtac cccaaaactc 3840
 caatttgcgtc ctggcccccag ccccttagaa ctgttgcgtt tccaggactg aacctctc 3900
 cccagaggcc tcagagggtga gcaggcttag caggggccag gcagtcctgc ggccactgac 3960
 gtcggctgg gagggtggac agggcccccgg ctcagagcca tctgtggtc tgggttagac 4020
 ccatccctcc tggctctgg gcctcagggt cctcctctgc agggtgaggt tgataaccac 4080
 cgcgctgcct ccctcccccac acagaacaacg gaaaacggcc ggggtggctg tcgtggctg 4140
 aattcaggga ggtatctcggt gatgaggatg gtcagtcca gaagagccaa gcaactgtg 4200
 tccctgaaat agggtgagcg gcagagactg gcccacccac acccacaccg cactgcgggg 4260
 ccctgtgtg caacgggtcc acacatcatc agccagggtgc cggcctgc tttgtcttc 4320
 aacctcgctg ccacagccac tgagttata cgagatttt gattttgtgc tttttgtcat 4380
 ttttatgaat caacatctgt caagcatctt ggccttaat gtttaggcgt gtgtgtgt 4440
 gtgtgtgtgt gtgtgtctca ctctgtcgcc cagactagag tgcaatgggg caatctctgt 4500
 tcactgcaac ctccacccctc caggttcaag cgattctctt ccctcagcct cccaaatgac 4560
 tgggattaca ggcggccgc accatgcgtg gctaattttt gatatttttag tagagacagg 4620
 gaacgtttag gcttaagtgt atttggagct gttggctt ctaggctaa acacctagag 4680
 atcagctggc cgctggcagc agcatgtgtt gtggacaca tgagctgtat ctatccca 4740
 ctcccatctt gccccatccc catgctccc ctccctctgg ctggcatcc ccactccctg 4800
 gcacccagggttggccata catccctccc atagcctcta ttcccatcg gtcctccac 4860
 aagggtttt tcctctgtt acccttgcac aagtcttatac accactcaga tgctacttcc 4920
 cctcaggatc cagcaactgtat gccccttccatgaaagcc cccagatccc tgccctgc 4980
 atgtccccc ctgtctccca cacagtggcc cagcccaactg accgcctgtc ccacacagcc 5040
 taccatgtgg tccagcaatt gctatctgtc actgttattat ctctcccgca ccctgtgacc 5100
 tcctctggc caggagatctt tccgtgttca tccctttagt aagaaggtgt atggcctgg 5160
 ggcgtggct caccgtgtat atcccaatc ttggggaggc caaggcagggt ggatcacctg 5220
 aggtcaggag ttccgagatca gcttggccaa catggtaaaa tccatctt actaaaaaaaa 5280
 atacaaaaat tagccggccg tgggtgtccgg cgcctgtat cccagctact cgggaggctg 5340
 aggccaggaga atcgcttgc gcccaggatgtt agaggatcata gtcgttgcgactgtgcac 5400
 tgcactccag cctggccaaatc aggtgaaatc tctgtctca aaaaaaaaaa aaaaaaaaaa 5460
 aaagaagcg cagcagcgc agcagcagca gctgcacaca tggggaaaata agagcaatca 5520
 gaccctcttca atttgcattt ccataggaag tttagggagat tactgttcca gcaactggct 5580
 cattccggaa aacagccgca gaagaaatga gaggcttctc aagtccgtca tttttttttt 5640

tttttttt gaaatatata atcaggcaat ctcccaagcc agaatggctc agagagactc 5700
 ccaagtctta catttcaag taaaattca ggcagtgaaa ctccagagca cacaactggg 5760
 ccatggctgt aaaccagcta cggtctgaat ctcagatcca tgaaaacaca gaggcagcct 5820
 caccactccg caggctgtga acaggacgga cgctggcta ctggccatct ccctgcctgc 5880
 ctgttgcattcc actttcaaca ggcctgagtc ttctctgac tcaccatca ggtggacccc 5940
 ctgcagcagg agtgttcccg gccacagaaa gccgattccc tcatctcatc actcaatacc 6000
 tatctactga gtctccctcc tcagagtcca gccctgcgtc gggggcacag agaggagtca 6060
 gctccccgtg gagacagaaa ccaggtactg gctggggcggg gggactgcag gagggggcag 6120
 ggtgctgaag aggggtccat ggcagttgtg gatggatgag ctgagctgg gaagggaaaca 6180
 gaggtggctc agggctgtga gacagacagg gaagcctaga gcagctgaag aacgggaagt 6240
 cagtggacca ggaagacagg aggccgggtt agaaccagat ggccgtcagc cttactcca 6300
 aggaattgaa agccaacatg gttcatgaa tgatgtttt gaagctggg gggccgggtg 6360
 cttagtctcc ctttctgaag ataacacccc gtcaggacc gttcacatg cggacagccc 6420
 tggcagcact taaaaggcac cagaagcgaa tgagaagcag aatcggtat gctgtctgagc 6480
 acgtgcacaa agcagacacg gttctcagga ctcgcattcc ttcaccact tcattccac 6540
 aacatccctaa agagggaaat acttcacccgta tgaggagacc gaggcacaga gcagttatgg 6600
 cacgtccccg aggtcacaca gtcagagatg gaaccggac ttgacccag gcaggctggc 6660
 tccagagcat ctccagtgtc ccaccacat gctggaggc tggggctt tctcagcagt 6720
 tcaactcaga cccagactcc acagcaggc ctgggggcca atcagaatag gccctctcac 6780
 ccaagggctc aggcataaaga cctgaacctg aagcacctct cgctcaaaca agcagctcag 6840
 gaacccaaa atgcagtgtg aataagaccc aggtcatacc catgcctggc tgcctccagg 6900
 atgtctatgg tctccaagaa taagctaaga ataagaatag gggctgtcg ttaggcagcc 6960
 cagcaagaga gcacctgtta gccggagcct cccatgaacc ttceggcaat gtttttgc 7020
 agcggcggtt cctttccctc tgacgggtcc cagtccaaacc aaaccagcc aacaggaagc 7080
 agcctggatt gtgttatcat tcagtcaatc gacaaacgct gaatctatac tgcacactct 7140
 gcacctgtt ccctactgga aaccaagctt ctccaggga acgggggtggg acaggaactg 7200
 aggggtcatg tggccctggc tggggccggg aacagggtca ggacagccct ccactgcccag 7260
 gcctggagtg taccctgcac ctaagccgtt ggtgtatgcac ctgtcaccgc aaaacattct 7320
 acagtatca ggggaacttc taccgttatttggcaaaacg agtgggtggct gtgtatgg 7380
 gatgggtatttaca caagactcac aaactcatca atctggacta aaggacaatc 7440
 ctaaaaagaat cgattctgcc cacagagacg tggtatgatc tgatgtttc tcaactccaa 7500
 agccagaaca gatactgcct accctctggg ggaaggatcc ttctctccc caatgcaagt 7560
 gagacaacat ggggtgtggc tctctagaac tagtggatcc cccgggaccg gtccccccac 7620
 tcacctgcca gtaagcagtg ggttctctag ttagccagag agctctgctt atatagaccc 7680
 cccaccgtac acgccttaccg cccatttgcg tcaatggggc ggagttgtta cgacattttg 7740
 gaaagttcccg ttgatattgg tgccaaaaca aactccatt gacgtcaatg ggggtggagac 7800
 ttggaaatcc cctgtgatca aaccgttatac cacgcccatt gatgtactgc caaaaccgca 7860
 tcaccatggt aatagcgatg actaatacgt agatgtactg ccaagtagga aagtccata 7920
 aggtcatgta ctgggcataa tgccaggcgg gccatttacc gtcattgacg tcaatagggg 7980
 gctgtacttgg catatgatac acttgatgtt ctgccaagtg ggcagttac cgtaaatagt 8040
 ccaccatttgc acgtcaatgg aaagtccctt ttggcggttac tatggaaaca tacgtcatta 8100
 ttgacgtcaa tggccgggg tgcgttggcg gtcagccagg cggggcattt accgtaaattt 8160
 atgttaacgcg gaactccata tatgggttat gactaatga ccccgtaattt gattactatt 8220
 aataactagt caataatcaa tgtcaacgcg cgggctgcag gaattctacc gggtagggg 8280
 ggcgcctttc ccaaggcagt ctggagcatg cgcttttagca gccccgctgg cacttggcgc 8340
 atcacaatgt gcctctggcc tcgcacatccatccca cccgttagcgc caaccggctc 8400
 ccttctttgg tggcccttgc ggcacccattt ctactctcc cctagtcagg aagttccccc 8460
 cccggccgcgca gctcgctcg tgcaggacgt gacaaatggg agtagcacgt ctcactagtc 8520
 tcgtgcagat ggacaaggcac cgctgagcaaa tgaaaggcggg taggccttgc gggcagcggc 8580
 caatagcaggc ttggctctgg gtcagaggc tggaagggg tgggtccggg 8640
 ggcgggctca gggccgggtt cagggccggg gccggccgcg aggtccctcc gacccggcat 8700
 tctgcacgct tcaaaaagcgc acgtctccg cgctgttctc ctcttcctca tctccgggcc 8760
 ttcgacccatc atgaaaaaaatc ctgaaactcac cgcgacgtct gtcgagaagt ttctgatcga 8820
 aaagttcgac agcgtctccg acctgtatgca gctctccggag ggcgaagaat ctcgtgctt 8880
 cagttcgat gtggatatgt cctgcgggtt aatagctcg ccgatggttt 8940
 ctacaaagat cggtatgttt atcggactt tgcattccggc ggcgtcccgaa ttccggaaagt 9000

gcttgcatt gggaaattca gcgagagcct gacctattgc atctccgcgtgtcacgttg caagacctgc ctgaaaccga actgcccgtcttctgcagccggcgatggcatggat g 9131

```
<210> 84
<211> 1121
<212> DNA
<213> Artificial Sequence
```

<220>
<223> Description of Artificial Sequence: Synthetic
EYPFP sequence

<400> 84
cttgtacagc tcgtccatgc cgagagtgtat cccggcgccg gtcacgaact ccagcaggac 60
catgtatcg cgcttctcggt tggggctttt gtcaggcg gactggtagc tcaggtatgt 120
gttgcgggc agcagcacgg ggccgtcgcc gatgggggtg ttctgtgtt agtggtcggc 180
gagctgcacg ctggccgtctt cgatgtgtg gccgatctt aagttcacct tgatgccgtt 240
cttctgttg tcggccatga tatagacgtt gtggctgtt tagtttactt ccagcttgtt 300
ccccaggatg ttggccgtctt ccttgaagtc gatgcccttc agctcgatgc ggttcaccag 360
ggtgtcgccc tcgaacttca cctcggcgcg ggtcttgttag ttggcgctgtt cttgaagaa 420
gatggtgdcg tcctggacgt agccttcggg catggcgac ttgaagaagt cgtgtgtt 480
catgtgttcg gggtagcgccc cgaagactg caggccgttag ccgaagggtgg tcacgggggt 540
gggcccaggc acgggcagct tgccgggtt gcatgttacat ttccagggtca gtttgcgtt 600
ggtgcatacg ccctcgccct cggcgacat gctgaacttgg tggccgttta cgtcgccgtc 660
cagtcgacc aggtatggca ccaccccggt gaacagcttc tggcccttgc tcaccatgtt 720
ggcggccgtt ggatcttct agagtgcacg atcgaagtta gcttgattt acagttggctt 780
ggggttgcgc cgccgggttt tataggaagc cacagcggcc actcgagcca taaaaggcaaa 840
cttttaggaac ggccgggggtt gattggattc gagtcgttta ttcacccggcc ttgcccaca 900
gtgcagcatt ttttacccc ctctccctc ctttgcggg ggaaaaaaaaaaaaaaa 960
aaaaaaaaaggag aagagaaaaaa aagcgagcga gagagaaaagc gagattgagg aagaggatgaa 1020
agagtttggc gatgggtgtt gttccgttag gcccagatgg acaagaatag ccccccgcctt 1080
tgccggacagt atcccatca gtgactcaga tcagatcaag c 1121

```
<210> 85
<211> 10557
<212> DNA
<213> Artificial Sequence
```

<220>
<223> Description of Artificial Sequence: Synthetic
PKI-CMV-SD-Vanilloid-YFP sequence

```

<400> 85
tttttacccc ctctccccctc cttttgcggg gaaaaaaaaaaa aaaaaaaaaaa aaaaaaaggag 60
aagagaaaaaa aaggcgagcga gagagaaaagc gagattgagg aagaggatga agagttggc 120
gatgggtgct ggttccgttag gccccagatgg acaagaatag ccccccgcct tgccgacagt 180
atcccatc a gtgactcaga tcagatcaag ccggccgctt gcccaagggcc ctgtgaagca 240
gggtacccca aaactccaaac ttgctgctgc ccccaagcccc ctagaactga agtgtttcca 300
ggactgaacc tctcacccca gaggccctcag aggtgagcag gcctagcagg gccccaggcag 360
tcctgcggcc actgacgtcg gcctgggagg gtggacaggc ccccgctca gqccatctg 420
tggttctggg ttagacccat ccctccctcg tctgggcct caggtgcctc ctctgcaggg 480
tgaggttgc aaccaccgcg ctgcctccct ccccacacag aacaaggaaa acggccgggg 540
tggctgtcg gggctgaatt cagggaggat ctgggtgatg aggtatggctc agtccagaaag 600
agccaagcag actgtgtccc tgaaataggg tgacggcag agactggccc acccacaccc 660
acaccqcact qcqqqqccct qctgtqcaac qqgtccacac atcatcagcc aggtgccccgg 720

```

cctgcctttt	ctcttcAAC	tcgctGCCAC	agccactGAG	cttatacGAG	atTTTGGATT	780
tttgctttt	tgtcattttt	atgaatcaac	atctgtcaag	catccTGGCC	ttGAATGTT	840
aggcggtgt	gtgtgtgtgt	gtgtgtgtgt	gtctcactct	gtcgcccaga	ctagagtGCA	900
atggggcaat	ctctgttCAC	tgcaacCTCC	acctcccAGG	ttcaagCGAT	tctccTCCCT	960
cagcccccca	agtagctGGG	attacaggCG	cccgCCACCA	tcgctGGCTA	atgtttGTAT	1020
tttttagata	gacagggAAC	gtttaggCTT	taagtGATT	ggagctGGTT	gtcttCTAG	1080
gcctaaacac	ctagagatca	gctggccgCT	ggcagcAGCA	tgttagttG	gacacatGAG	1140
ctgatgttat	tttccactCC	catcctgCCC	catccccATG	cttccCTCC	ctctggCTGG	1200
gcatccccac	tccctggcac	ccagggGTGG	ccatacATC	cctccatAG	cctctattCC	1260
catcggtct	ccacctaagg	gtttttCTCT	ctgtctACCC	ttgcacaAGT	cctatCACCA	1320
ctcagatgt	acttcccCTC	aggatCCAGC	actgatGCCG	cctccCTCAT	gaagCCCCCA	1380
gatccctgCC	ctgctcatGC	tccccgtGT	cctccacACA	gtggcccAGC	ccactgACCG	1440
cctgtccccac	acagcctACC	atgtgttCCA	gcaattGCTA	tctgtcACTG	atttatCTCT	1500
cccgcacccct	gtgacctCTC	cctgggcAGG	gaccTTTCTC	ggtccatCTC	tttatGAAGA	1560
agctgtactg	gcctgggCGC	gtggctCACG	cctgtaatCC	cagcacTTG	ggaggcCAAG	1620
gcaggggtat	cacctgAGGT	caggagtTCG	agaccAGCTC	ggccaaACATG	gtgaardATCC	1680
atctctacta	aaaaaaaATAC	aaaaatAGC	cggggctGTTG	gtcggggcCC	tgtatCCCA	1740
gctactcGGG	aggctgAGGC	aggagaATCG	cttgaacCCA	ggatgtAGAG	tttacAGTGA	1800
gctgagactg	tgccactGCA	ctccagCCTG	ggcaacAGAG	tgaaactCTG	tctcaaaaaAA	1860
aaaaaaaaaa	aaaaaaaaAG	aagcagCAGC	agcagcAGCA	gcagcAGCTG	cacacatGGG	1920
gaaataaAGAG	caatcAGACC	tttctaATTG	ccatttCCAT	aggaagtTAG	ggagattACT	1980
gttccagcag	tggctacATT	ccggaaaACA	gcccgcAGAAG	aatgagAGG	tttctcaAGT	2040
cctgcatttt	cttttttttt	tttttGAAA	tatataatCA	ggcaatCTCC	caagCCAGAA	2100
tggctcagag	agactCCAA	gtcctacATT	tcaagttAA	aattcaggCA	gtgaaactCC	2160
agagcacaca	actgggCCAT	ggctgtAAAC	cagctacGTT	ctgaatCTCA	gatccatGAA	2220
aacacagagg	cagcctACC	actccgcAGG	ctgtgaACAG	gacggacGGC	tggctactGG	2280
ccatctccct	gcctgcCTG	tgtccactt	tcaacacGGCC	ttagtCTTC	tctgactCAC	2340
ccatcaggTG	gaccccCTG	agcaggAGTG	ttcccgGCCA	cagaaAGCCG	attccCTCAT	2400
ctcatcactc	aatacCTATC	tactgagtCC	toctcCTCA	agtccAGCCC	tgcgctGGGG	2460
gcacagagag	gagtcaGTC	cccgctGAGA	cagaaACCA	gtactggCTG	ggcggaggGA	2520
ctgcaggagg	gggcaggGTG	ctgaagAGGG	gtccatGGCA	gttgggtATG	gtatgAGCTG	2580
gcctgggaag	ggaacAGAGG	tggctAGGG	tctggagACA	gacagggAAAG	cctagAGCAG	2640
ctgaagaACG	ggaagtCAGT	gggacAGGAA	gacaggAGGC	cgggttagAA	ccagatGGCC	2700
gtcagcTTA	actccaAGGA	attgaaAGCC	aacatGGTTG	catggatGAT	gtttggAAAG	2760
cctggggggc	cgggtgCTAG	tcttccTTT	ctgaagatAA	cacccGTCA	aggaccGTC	2820
acagtgcGGA	cagccCTGGC	agcacttAAA	aggcaccAGA	agcgaatGAG	aagcagaATC	2880
gttgtatgcgt	ctgagcACGT	gcacaaAGCA	gacacGGTC	tcaggactCG	cattcCTCA	2940
cccacttcat	tcccacaACA	tccctaAGAG	gaaaatACTT	caccGATGAG	gagaccGAGG	3000
cacagagcag	ttatggCACG	tccccgAGGT	cacacAGTC	gagatGGAAC	ccggacttGA	3060
ccccaggcag	gctggcTCC	gagcatCTCC	agtgcTCCAC	cacGATGCTG	gaggGCTGG	3120
gcttttCTC	agcagtTCAA	ctcagacCCCA	gactCCACAG	caggGCTGG	gggcaatCA	3180
gaataggccc	tctcacCCAA	ggggctAGGC	ataagacCTG	aacctGAAAGC	acctctCGCT	3240
caaacaAGCA	gctcaggAAC	ccaaaATGC	agtgtGAATA	agaccCAGGT	cataccCATG	3300
cctggctgtc	ccaggGATGT	ctatggCTC	caagaataAA	ctaagaataAA	gaataggGGC	3360
tgctcgTTAG	gcagccccAGC	aagagAGC	ctgttagCCG	gagccTCCC	tgaacCTTC	3420
ggcaatgttc	tttgtcAGCG	gggttgcTTT	ttctctGAC	gggtcccAGT	ccaaacAAAC	3480
ccagccaaaca	ggaaggcAGGC	tggattgtGC	tatcattCA	tcagtcGACA	aacgtGTAAT	3540
ctatactgcA	cactctGCAC	ctgatTTCT	actggAAACC	aagcttCTTC	caggGAACGG	3600
ggggcggacag	gaactgAGGG	gtcatgtGCC	ctgggCTGGG	ccgggAAACA	gggtcaggAC	3660
agccctccac	tgccaggCCT	ggagtgtACC	ctgcacCTAA	ggccgtggTG	atgcacCTGT	3720
caccgcaaaa	cattctACAG	tgatcaGGGG	aacttCTACC	gtgattTAGG	caaacGAGT	3780
ttggctgtgt	gatggagatG	gtgattCTGG	tttacacaAG	actcacAAAC	tcatcaatCT	3840
ggactaaagg	acaatcCTAA	aagaatCGAT	tctgcccACA	gagacGTTG	atgatCTGAG	3900
tctttctcaa	ctccaaAGCC	agaacAGATA	ctgcctACCC	tctggggAA	ggatccCTTC	3960
tctccccaat	gcaagtGAGA	caacatGGGG	tgtggctCTC	tagaactAGT	ggatccccCG	4020
ggaccggTCC	ccccactCAC	ctgcccAGTA	gcagtGGGTT	ctctagTTAG	ccagagAGCT	4080
ctgcttatat	agacctCCCA	cgtaCACGC	ctaccGCCA	tttgcgtCAA	tgggGGGGAG	4140
ttqttacgac	attttggAAA	gtcccgttGA	tttgggtGCG	aaaacAAACT	ccattGACG	4200

tcaatgggt ggagacttgg aaatccccgt gagtcaaacc gctatccacg cccattgatg 4260
 tactgccaaa accgcatacac catggtaata gcgatgacta atacgttagat gtactgccaa 4320
 gtaggaaagt cccataaggc catgtactgg gcataatgcc aggcgggcca ttaccgtca 4380
 ttgacgtcaa tagggggcgt acttggata tgatacacactt gatgtactgc caagtggca 4440
 gtttaccgt aatagtccac ccattgacgt caatggaaag tccctattgg cgtaactatg 4500
 ggaacatacg tcattattga cgtcaatggg cgggggtcg tggcgggtca gccaggcggg 4560
 ccatttaccg taagttatgt aacgcggAAC tccatatacg ggctatgaac taatgacccc 4620
 gtaattgatt actattaacta actagtcaat aatcaatgtc aacgcggggg ctgcaggaat 4680
 tctaccgggt agggggggcg ctttcccaa ggcagtctgg agcatgcgt ttagcagccc 4740
 cgctggcact tggcgcata caagtggct ctggcctcg acacattcca catccaccgg 4800
 tagcgcacac cggtccctt ctgggtggc cccttcgac caccttctac tcctccctt 4860
 gtcaggaagt tccccccgc cccgcagctc gcgtcggtca ggacgtgaca aatggaaatg 4920
 gcacgtctca ctatgtctcg gcagatggac aagcaccgc gagcaatgga agcgggtagg 4980
 ccttggggc agcggccaaat agcagcttgg ctccctcgct ttctggctc agaggctggg 5040
 aagggggtggg tccggggggcg ggctcagggg cgggctcagg ggcggggggcg ggcgcaaggt 5100
 cctccggacc cggcattctg cacgcttcaa aagcgcacgt ctggcgcgt gttctctt 5160
 tcctcatctc cgggccttcg acctgcataa aaaagcctga actcaccgcg acgtctgtcg 5220
 agaagtttct gatcggaaatg ttcgacagcg ttcggacactt gatgcagctc tcggagggcg 5280
 aagaatctcg tgcttcagc ttgcgtatgt gaggggcgtgg atatgtcctg cgggtaaata 5340
 gctgcggca tggtttctac aaagatcgatgtt atgtttatcg gcaacttgca tcggccgcgc 5400
 tcccgattcc ggaagtgtctt gacattgggg aattcagcga gagcctgacc tattgcatct 5460
 cccggcgtgc acagggtgtc acgttgcag aacctgcctga aaccgaactg cccgctgttc 5520
 tgcagccggc cgcggaggcc atggatgcga tcgctgcggc cgatcttagc cagacgagcg 5580
 ggttcggccc attcggaccg caaggaatcg gtcaatacac tacatggcgt gatttcatat 5640
 ggcgcattgc tgatccccat gtgtatcact ggcaaaactgt gatggacgac accgtcagtg 5700
 cgtccgtgc gcaggctctc gatgagctga tgctttggc cgaggactgc cccgaagtcc 5760
 ggcacccgt gcaacggat ttcggctcca acaatgtcct gacggacataa ggccgcataa 5820
 cagcggcat tgactggagc gaggcgatgt tcgggggattt ccaatacagag tgcggcaaca 5880
 tttttctg gaggccgtgg ttggctgtt tgaggcagca gacgcgttac ttgcggcgga 5940
 ggcattccgga gtttgcagga tcgcccggc tccggggcgtatg tgcgtccgc attggcttt 6000
 accaactcta tcagagctt gttgacggca aattcgtatg tgcagcttgg ggcgggggtc 6060
 gatgcacgc aatcgtccga tccggagccg ggactgtcgg ggcgtacacaa atcgcccgca 6120
 gaagcgcggc cgtctggacc gatggctgtg tagaagttact cgccgatagt ggaaaccgac 6180
 gccccagcac tcgtccgagg gcaaaaggaaat gagctcgctg atcaggctcg actgtgcctt 6240
 ctagttgcca gcatctgtt gtttgcctt ccccccgtgcc ttccctattgt ctgagtaggt 6300
 gtcattctat tctgggggggtt ggggtggggc aggacagcaaa gggggaggat tgggaagaca 6360
 atagcaggca tgctggggat gcccgtggct ctatggcttc tgaggcggaa agaaccagct 6420
 gggggatatac aagcttatacg ataccgtcgatc cctcgagggg gggccggta cccagcttt 6480
 gttccctta gtgagggtt attgcgcgt tggcgtatc atggcatacg ctgtttcctg 6540
 tgtgaaattt ttatccgtc acaattccac acaacatacg agccggaaagc ataaagtgt 6600
 aagctgggg tgcctaatacg gtgagctaac tcacattaat tgcgttgcgc tcactgcccc 6660
 cttccagtc gggaaacactg tcgtccgacg tgcatatgtt aatcgccaa cgccgggg 6720
 gaggcgggtt ggttattggg cgctctccg ctccctcgct cactgactcg ctgcgtctgg 6780
 tcgttcggct gcccggcgtc gtatcgactc actcaaaggc ggtaatacgg ttatccacag 6840
 aatcaggggta taacgcaggaa aagaacatgt gaggaaaaagg ccagaaaaag ggcaggaaacc 6900
 gtaaaaaggc cgcgttgcgt ggcgttttcc ataggctccg ccccccgtac gaggcatcaca 6960
 aaaatcgacg ctcagaatcg aggtggcgaa acccgacagg actataaaaga taccaggcg 7020
 ttccccctgg aagctccctc gtgcgtctc ctgttccgcg cctggccgtt accggataacc 7080
 tgcctccctt tctccctcg ggaaggcgatgg cgcttctca tagctcagcg tggtaggtatc 7140
 tcagttcggt ttaggtcgatc cgctccaaacg tggctgtgtt gcaacgaaacc cccgttgcgt 7200
 ccgaccgctg cgcccttatcc ggttaactatc gtcttgatc caacccggta agacacgact 7260
 tategcactt ggcaggcgtc actggtaaca ggatttagcag agcgaggtat ttagggcggt 7320
 ctacagagtt ctggtaactt gggcttacac acggctacac tagaaggaca gtatggta 7380
 tctgcgtct gctggtaacca gttacccgtt gaaaaaggt tggtagctt tgcgttgcgc 7440
 aacaaaccac cgctggtagc ggtgggtttt ttgttgcac gcaacgacatt acgcgcagaa 7500
 aaaaaggatc tcaagaagat ctttgcattt ttgttacggg gtctgcgtc cagtgaaacg 7560
 aaaactcactg ttaaggggatt ttgttgcataa gattatcaaa aaggatctt accttagatcc 7620
 ttttaaatta aaaatgaagt tttaaatcaatcataaacttggctct 7680

acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta ttcgttcat 7740
 ccatagttgc ctgactcccc gtcgttaga taactacgt acgggagggc ttaccatctg 7800
 gccccagtgc tgcaatgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa 7860
 taaaccagcc agccggaagg gccgagcgca gaagtggtcc tgcaactta tccgcctcca 7920
 tccagtttat taattgttgc cgggaagcta gagtaagtag ttgcgcagtt aatagttgc 7980
 gcaacgttgt tgccattgtc acaggcatcg tgggtcacg ctcgtcggtt ggtatggctt 8040
 cattcagtc cggccccaa cgatcaaggc gagttacatg atccccatg ttgtgcaaaa 8100
 aaggcggttag ctccctcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 8160
 cactcatgtt tatggcagca ctgcataatt ctcttactgt catgcattcc gtaagatgtc 8220
 tttctgtgac tggtagtac tcaaccaagt cattctgaga atagttatg cggcgaccga 8280
 ttgtcttgc cccggcgtca atacggata ataccgcgc acatagcaga actttaaaag 8340
 tgctcatcat tggaaaacgt tttcggggc gaaaacttc aaggatctt ccgcgttga 8400
 gatccagttc gatgtAACCC actcgtgcac ccaactgatc ttcagcatct ttactttca 8460
 ccagcggttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataagg 8520
 cgacacggaa atgttgaata ctcatactt tccttttca atattattga agcatttattc 8580
 agggttattt tctcatgago ggatacatat ttgaatgtat tttagaaaaat aaacaaatag 8640
 gggttcccgcg cacatttccc cgaaaagtgc cacctaaatt gtaagcgtta atattttgtt 8700
 aaaattcgcg taaaattttt gttaaatcag ctcatttttt aaccaatagg cggaaatcg 8760
 caaaatccct tataaatcaa aagaatagac cgagataggg tttagtggg ttccagttt 8820
 gaacaagagt ccactattaa agaacgtgga ctccaaacgtc aaagggcga aaaccgtcta 8880
 tcagggcgat ggcccactac gtgaaccatc accctaatac agtttttgg ggtcgaggtg 8940
 ccgtaaagca ctaaatcgga accctaaagg gagccccga tttagagctt gacggggaaa 9000
 gccggcgaac gtggcgagaa aggaaggaa gaaagcgaag ggagcggcg cttagggcgct 9060
 ggcaagtgtt gcggtcacgc tgcgcttaac caccacaccc gccgcgtta atgcgcgcgt 9120
 acagggcgcg tccattcgc cattcaggct gcgcactgt tgggaaggc gatcggtgc 9180
 ggccttccg ctattacgca agctggcgaa agggggatgt gctgcaaggc gattaagttt 9240
 ggttaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg agcgcgcgta 9300
 atacgactca ctataggcg aattggagct ccaccgcggg ggcggccgca acagacatga 9360
 taagatacat tgatgatgtt ggacaaacca caactagaat gcaatggaaa aaatgtttt 9420
 ttgtgaaat ttgtgatgtt attgtttat ttgttaaccat tataagctc aataaaacaag 9480
 ttgggggtggg cgaagaactc cagcatgaga tccccgcgtt ggaggatcat ccagccggcg 9540
 tccccggaaaaa cgattccgaa gcccaccc ttcatagaagg cggcgggtgga atcgaatct 9600
 cgtatggca ggttggcggt cgttggctg gtcatttca agctttactt gtacagctcg 9660
 tccatgccga gaggatccc ggcggcggtc acgaactcca gcaggaccat gtgatcgccg 9720
 ttctcggtgg ggtctttgtc cagggcgac ttgttagctca ggttagtggg gtcgggcagc 9780
 agcacggggc cgtcgccgat ggggggttgc ttgtggtagt ggtcggcgag ctgcacgctg 9840
 ccgtcctcga ttgtgtggcg gatcttgaag ttacacctga tgccgttctt ctgcttgc 9900
 gccatgatat agacgttgc gctgtttagt ttgtactcca gcttgtgccc caggatgtt 9960
 ccgtcctcct tgaagtgc gcccctcgc tcgatgcggg tcaccagggt gtcgcctcg 10020
 aacttcaccc cggcgccgggt cttgttagtt ccgtcgctt tgaagaagat ggtgcgcctc 10080
 tggacgttagc ttccggcat ggcggacttg aagaagtctg gctgcttcat gtggtcgggg 10140
 tagcggcgca agcaactgcag gccgtcgcc aagggtgtca cgagggtggg ccagggcact 10200
 ggcagcttgc cgggtgtca gatgaacttc agggtcagct tgccgttagt ggcatcgccc 10260
 tcgcctcgc cggacacgct gaacttgtgg ccgttacgt cggcgtccag ctcgaccagg 10320
 atgggcacca ccccggtgaa cagctcctcg cccttgctca ccatggtggc gaccgggtgg 10380
 tccttctaga gtgcacgatc gaagtttagt tgatttgaca gtggctgggg gttgcgcgc 10440
 cgggttttat aggaagccac agcggccact cgagccataa aaggcaactt taggaacggc 10500
 ggggggtgat tggattcgcg tgcgttattc accggccttgc ccgcacagtgc cagcattt 10557

<210> 86
 <211> 3799
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic nucleotide sequence for Vanilloid region of homology

<400> 86

ttcaagcgat ttcctgcct cagcctccc agtagctggg attacaggcg cccgacacca 60
 cgccccgcta attttgtat ttttttagt agagatggg tttcacatg ttggccaggc 120
 tggctcgaa ctccctgacct caggtgatcc acctgccttgcctccaa gtgctggat 180
 tacaggcgtg agccacgcgc ccaggccagt acagcttct cataaggaga tggaccagga 240
 aaggtcctg cccaggagga ggtcacaggg tgccggagaa ataatacgat gacagatagc 300
 aattgtcgaa ccacatggta ggctgtgtgg gacaggcggt cagtggctg gcccacttg 360
 tggaggacag cggggagcat gaggcaggca gggatctggg ggcttcatgg aggaggcggc 420
 atcagtctg gatcctgagg ggaagtgca ttctgtgtt gataggactt gtgcaagggt 480
 agacagagga aaaagccctt aggtggagac cogatggaa tagaggttat gggagggatg 540
 tatggccac ccctgggtgc cagggagtg ggtatggccag ccagagggag gggaaagcatg 600
 gggatggggc aggatggag tggaaaatag catcagctca tttgtccaac tacacatgt 660
 gctgccagcg gccagctgat ctctaggtt tttaggcctt aagaccaacc agctccaaat 720
 cacttaaagc ctaaacgttc cctgtctcta ctaaaaaatac aaacattagc cacgcatgt 780
 ggcgggcgc ttaatccca gctacttggg aggctgaggg aggagaatcg ttgaacctg 840
 ggagggtggag gttgcagtga acagagattt cccattgc cttctagtc ggcgacagag 900
 ttagacacac acacacacac acacacacac acgcctaaac attcaaggcc aggtgctt 960
 acagatgtt attcataaaaa atgacaaaaaa gcacaaaatc caaaatctg tataagctca 1020
 gtggctgtgg cagcgagggtt gaagagcaaa ggcaggccgg gcacctggct gatgtgtt 1080
 ggaccgttg cacagcaggg ccccgactg cgggtgtgggt gtgggtggc cagtctctg 1140
 cgctcaccctt attccaggga cacagtcgtc ttggctcttc tggactgagc catcctcatc 1200
 accgagatcc tccctgaattt cagcccacga cagccacccc ggccgtttt cttgttctgt 1260
 gtggggaggg aggcaagcgcg gtggttatca acctcaccctt gcagaggagg cacctgaggc 1320
 ccagagacga ggagggtatgg gtctaaacca gaaccacaga tggctctgag cccggggcct 1380
 gtccaccctc ccaggccgac gtcagtggcc gcaggactgc ctggccctg ctaggcctgc 1440
 tcacctctga ggctctggg gtgagagggtt cagtcctgg aacacttc tttctaggggg 1500
 ctggggcgcag cagcaagttt gatgtttggg gtaccctgtc tcacaggggcc ttggcaagc 1560
 ggcgcctcag atctagagag ccacacccca ttgtgtctca cttgcatttgg ggagagaagg 1620
 gatcctccccc ccagggggta ggcagtatct ttctgtggct tggagttgag aaagacttag 1680
 atcataccac gtctctgtgg gcagaatcgat ttcttttagt attgtctttt agtccagatt 1740
 gatgagttt tgagtctgtt gtaaaccaga atcaccatct ccatcacaca gccaacactc 1800
 gtttgcctaa atcacggtag aagttccctt gatcaactgtt gatattttgcgtgacagg 1860
 tgcacatcca cgggctttagg tgcagggtac actccaggcc tggcagtggaa gggctgtcct 1920
 gaccctgttc cccggcccag cccaggccac atgacccctc agttctgtc ccaccccggt 1980
 ccctggaaaga agcttggttt ccagtagggaa atcagggtca gagtgtgcag tataattca 2040
 gctttgtcg actgactgaa tgatagcaca atccaggctg cttectgtt gctgggtttt 2100
 gttggactgg gacccgtcag aggaaaaaggc aacgcgcgtg acaaagaaca ttgcccggaa 2160
 gttcatggga ggctccggct aacagggtct ctctgtctgg gctgcctaaac gaggcccc 2220
 tattcttattt cttagcttattt tcttggagac catagacatc cctggacacag ccaggcatgg 2280
 gtatgacactg ggtcttattt acactgcatt ttgggttcc tgagctgtt gtttgagcga 2340
 gaggtgttcc aggttcagggtt cttatgccta gccccttggg tgagaggccc tattctgatt 2400
 ggccccccagg ccctgtgtg gagtctgggt ctgagttgaa ctgctgagaa agagccccag 2460
 acctccagca tcgtgtgttgc gcactggaga tgcctgtggag ccagcgtcc tgggtcaag 2520
 tccgggttcc atctctgact gtgtgacccctt ggggacgtgc cataactgtt ctgtgcctcg 2580
 gtctccatcg cggtaagta tttcccttctt aggatgttg tggaaatgaa gtgggtgaag 2640
 gaatgcgagt cctgagaacc gtgtctgtt tgcacatgtt ctcagacgc tcaacgattc 2700
 tgcttcatcg tgcctctgg tgcctttaa gtgctgtggccag ggctgtccgc actgtgaacg 2760
 gtccttgacg ggggttttac ttccagaaagg gaagacttagc accccggccc ccaggcttcc 2820
 aaagcatcat ccatgcaacc atgttgcattt tcaatttctt ggagttaaagg ctgacggcca 2880
 tctggttcta accggccctc ctgtcttctt gttccactga ctcccgtt ttcagctgt 2940
 cttaggttcc ctgtctgttcc ccagacccttta gcccacccctt gttcccttcc caggctcagc 3000
 tcatccatca ccaactgcca tggacccctt ttccagccccc tggcccttcc tgcagtccct 3060
 ccggccagcc agtacctggt ttctgtctca gccccggagct gactcctctc tggccccc 3120
 ggcgagggtt ggactctgag gaggaggact cagtagatag gtattgagtg atgagatgag 3180
 ggaatcggtt ttctgtggcc gggaaacactc ctgctgcagg gggtcacactt gatgggttag 3240
 tcagagaaaag actcaggccctt gttgaaatgtt gatcaacacgg caggcaggga gatggccagt 3300
 agccagccgtt cctgtctgtt ccacagctgc ggagtggtt ggtgcctctt gtgttttcat 3360

102/102

ggatctgaga	ttcagaacgt	agctgggta	cagccatggc	ccagttgtgt	gctctggagt	3420
ttcactgcct	gaattttAAC	ttgaaaatgt	aggacttggg	agtctctctg	agccattctg	3480
gcttgggaga	ttgcctgatt	atatatTCa	aaaaaaaaaa	aaaagaaaaat	gcaggacttg	3540
agaaggctct	catttcttct	gcggctgttt	tccggaatgt	agccactgtct	ggaacagtaa	3600
tctccctaac	ttcctatgga	aatggcaatt	agaagggtct	gattgtcttt	atttccccat	3660
gtgtgcagct	gctgctgtcg	ctgctgctgc	tgcttcTTT	tttttttttt	tttttttttt	3720
tgagacagag	tttcactctg	ttgcccaaggc	tggagtgcag	tggcacagtc	ttagctcaCT	3780
gtAACCTCTA	catcctggg					3799



Creation date: 09-08-2004
Indexing Officer: SGIRON - SERGIO GIRON
Team: OIPEBackFileIndexing
Dossier: 10708420

Legal Date: 07-15-2004

No.	Doccode	Number of pages
1	PEFR	5
2	OATH	2
3	A.PE	1
4	CLM	5
5	REM	1
6	IDS	4
7	FOR	68
8	FOR	11
9	FOR	22
10	FOR	16
11	NPL	12
12	NPL	12

Total number of pages: 159

Remarks:

Order of re-scan issued on